Physical Mechanics

An Intermediate Text for Students of the Physical Sciences

BY

ROBERT BRUCE LINDSAY, PH.D.

Hazard Professor of Physics in Brown University

SECOND EDITION - SECOND PRINTING





D. VAN NOSTRAND COMPANY, Inc.
TORONTO NEW YORK LONDON



NEW YORK

- D. Van Nostrand Company, Inc., 250 Fourth Avenue, New York 3 TORONTO
- D. Van Nostrand Company (Canada), Ltd., 228 Bloor Street, Toronto LONDON

Macmillan & Company, Ltd., St. Martin's Street, London, W.C. 2

COPYRIGHT, 1933, 1950

BY

D. VAN NOSTRAND COMPANY, Inc.

Published simultaneously in Canada by

D. VAN NOSTRAND COMPANY (Canada) Ltd.

All Rights Reserved

This book or any parts thereof may not be reproduced in any form without written permission from the publishers.

> First Edition May 1933 Ten Reprintings Second Edition May 1950

Reprinted September 1951



PREFACE

The fundamental purpose of the revised version remains the same as that of the original edition published in 1933, namely, to serve as the basis of an intermediate course in mechanics for college students who have taken a substantial course in elementary general physics and college mathematics through calculus and introductory differential equations. The aim is still to stress the fundamental concepts and principles of mechanics and their use in all branches of physics. Certain additions, e.g., in atomic theory, relaxation phenomena, motion in resisting media, kinetic theory, oscillations, etc., reflect an increased concern with the applications.

An outstanding change has been made in the treatment of rigid bodies by the consolidation of the original two chapters on the statics and dynamics of rigid bodies, respectively, into a single chapter in which the dynamical approach is kept paramount and statics is brought in as a special case. This reorganization is not only a more logical carrying out of the original plan of the book but serves to de-emphasize somewhat the statics part which is not so important for physicists. In the development greater use has been made of vector notation, though as before the author has refrained from a detailed presentation of vector differential operators. The approach to rigid bodies is also made smoother and more logical by introducing earlier the chapter on the mechanics of collections of particles of which rigid bodies form a special case. The first part of this chapter has been largely rewritten with greater emphasis on the two-particle problem.

The fundamental concepts of impulse, momentum, work and energy are now introduced in Chapter I from an elementary point of view, thus remedying what several have considered a weakness in the first edition in which the idea of energy is first introduced in Chapter IV. Collision phenomena have been taken out of Chapter XI and introduced where they more naturally belong in Chapter VI. The material on advanced mechanics and wave mechanics has been placed in a separate chapter at the end of the

book.

The problem lists have been revised and more carefully graded. Numerous additions have been made, and answers to many prob-

lems are provided.

Occasion has been taken to correct the obvious errors and infelicities which occurred in the first edition, many of which have been brought to the attention of the author by those who have used the book. Several demonstrations have been materially simplified by putting more emphasis on the physical content than on algebraic manipulation.

The author wishes to express here his deep appreciation to all who have taken an interest in the book by forwarding criticisms and suggestions. He is particularly grateful to his wife for assist-

ance with the proof.

R. B. LINDSAY

Providence, R. I. February, 1950



CONTENTS

CHAPTER I

	THE ELEMENTAL CONCEPTS OF MECHANICS	
		PAGE
1.1.	What Is Mechanics?	1
1.2.	Fundamental Definitions	2
1.3.	Displacement and Velocity	6
1.4.	Linear and Angular Motion	11
1.5.	Acceleration	12
1.6.	Concept of Inertia and Definition of Mass	15
1.7.	Force and the Laws of Motion	17
1.8.	Units	19
1.9.	The Third Law of Motion. Combination of Masses	21
1.10.		24
1.11.	Résumé	29
	Problems	31
	CHAPTER II	
	RECTILINEAR MOTION OF A PARTICLE	
	ILECTIONEAR INITION OF A TARTICUE	
2.1.	Free Motion and the Uniform Field	33
2.2.	Motion in a Field Proportional to the First Power of the	
	Distance	37
2.3.	Motion in a Field Proportional to the Inverse Square of the	
	Distance	44
2.4.	Motion in a Field Proportional to the Inverse Cube of the	
	Distance	51
2.5.	Force Dependent on the Time. Impulsive Force	52
2.6.	Rectilinear Motion in a Resisting Medium	53
	Problems	55
	CHAPTER III	
	C	
	CURVILINEAR MOTION IN A PLANE	
3.1.	Components of Motion in a Plane	58
3.2.	Equations of Motion for a Particle in a Plane.	59
3.3.	Projectile Motion	60
3.4.	Projectile Motion in a Resisting Medium.	62
3.5.	Composition of Simple Harmonic Motions in a Plane	64
3.6.	Central Forces.	69

vii

		PAG
3.7	Square of the Distance	76 77 82
3.9 3.10	O. Motion in a Repulsive Inverse Square Force Field. Alpha Particle Deflection	87 90
	CHAPTER IV	
	ENERGY IN PARTICLE DYNAMICS	
4.1. 4.2. 4.3. 4.4.	Potential Energy and the Energy Concept. Conservative Forces Energy Relations in a Uniform Field Energy Relations in a Central Force Field Inverse Square Field. Electron Energies in the Bohr Theory	92 96 97 100
4.5.	Potential Calculations	103
4.6.	Further Considerations on the Potential. Gauss' Law and Laplace's Equation Problems	107 113
	CITY DAMED AT	
	CHAPTER V	
	STATICS OF A PARTICLE	
5.1. 5.2. 5.3.	General Equations of Equilibrium of a Particle	116 119 121
5.4.	Equilibrium of a Flexible String.	125
5.5. 5.6.	Equilibrium of a Particle on a Smooth Surface. Equilibrium of a Particle on a Rough Surface. Static Frica	127
5.7.	The Principle of Virtual Work	129 132
5.8.	Problems	35 38
	CHAPTER VI	
	Motion of a System of Particles	
6.1.	Equations of Motion. Center of Mass. Conservation of	
6.2.	Conservation of Momentum in Collisions	41
6.3.		43 45
6.4. 6.5.		±3 49
6.6.		53
		57

CONTENTS

ix PAGE Elementary Kinetic Theory of Gases. Equation of State of an Ideal Gas.... 159 Some Consequences of the Kinetic Theory. Specific Heats 6.8. of Gases..... 163 Elementary Kinetic Theory of the Viscosity of a Gas..... 166 6.10. Virial for Interaction Forces. Kinetic Theory of a Real Gas 171 Problems.... 172CHAPTER VII MECHANICS OF A RIGID BODY 7.1. Definition of a Rigid Body. Types of Motion..... 7.2. Rotation and Angular Velocity..... 176 7.3. Rotation about a Fixed Axis..... 178 7.4. Moment of Inertia Calculations.... 180 7.5. The Physical Pendulum.... 187 7.6. Plane Motion of a Rigid Body..... 191 7.7. General Equations of Motion of a Rigid Body..... 193 7.8. Center of Mass of a Rigid Body..... 196 Equilibrium of a Rigid Body. Center of Gravity...... 7.9. 202 7.10. Equilibrium under Coplanar Forces. Illustrations...... 209 7.11. Moving Axes..... 216 7.12. More About Moving Axes. Motion of a Particle on the Earth's Surface..... 219 7.13. Kinetic Energy of a Rigid Body..... 7.14. Euler's Equations of Motion..... 7.15. The Motion of a Top..... 229 Problems. 234 CHAPTER VIII CONSTRAINED MOTION Simple Types of Constraints. The Simple Pendulum..... 239 8.1. 8.2. Motion of a Particle on a Smooth Surface of Arbitrary Form 243 Constraints and the Principles of Mechanics. D'Alembert's 8.3. Principle—Dynamics Reduced to Statics..... 250 Gauss' Principle of Least Constraint..... 253 8.4. Problems. 257 CHAPTER IX OSCILLATIONS A Simple Problem in Vibration..... 2599.1. 9.2. Oscillations of a Dynamical System with One Degree of

Freedom—Dissipation.....

260

9.3 9.4 9.5 9.6 9.7 9.8	 Forced Oscillations of a Dissipative System. The Acoustic Resonator as an Illustration of Oscillator Motion. Electrical Oscillations. The Oscillator in Atomic Theory. 	273 cy * 281 · 284 · 285 · 289
	1100101110	. 294
	CHAPTER X	
	DEFORMABLE BODIES AND WAVE MOTION	
10.1. 10.2. 10.3. 10.4. 10.5. 10.6. 10.7.	Strain and Stress. Hooke's Law. Relations Among the Elastic Moduli. Elastic Limit. Fatigue and Heredity. Wave Motion. Transverse Waves in a String. Types of Elastic Waves in Solids. The Elastic Medium Theory of Light. Problems.	305 311 315 325 331
	CHAPTER XI	
	MECHANICS OF FLUIDS	
11.1. 11.2. 11.3. 11.4. 11.5. 11.6. 11.7. 11.8.	Fluids at Rest. Fundamental Principles of Hydrostatics. Principle of Archimedes—Stability of Floating Bodies The Equation of Continuity in Fluid Motion. The Equations of Motion of a Perfect Fluid. Steady Flow of a Liquid. Bernoulli's Theorem and Applications. Waves in Fluids. Viscous Fluids. Surface Phenomena. Capillarity. Problems.	338 341 345 349 352 358 365 370 382
	CHAPTER XII	
	ADVANCED MECHANICS	
12.1. 12.2. 12.3. 12.4.	Hamilton's Principle. Generalized Coordinates and Lagrange's Equations. Application of Lagrange's Equations. Wave Mechanics. Problems.	386 391 394 399 406
	Supplementary Problems. Answers to Selected Problems. Index.	409 429 435



CHAPTER I

THE ELEMENTAL CONCEPTS OF MECHANICS

1.1. What is Mechanics? Mechanics is the science of motion. Since motion is probably the most obvious of all physical phenomena, it is not surprising that the field of physics dealing with it has received careful attention by physicists from the very earliest times. The evolution of mechanics as we know it today has been a long protracted process involving the introduction and discarding of many different notions. The ideas of Galileo, Huygens and Newton, developed mainly during the seventeenth century, proved so successful in the description of the observed large-scale motions of bodies that it became only natural to attempt the explanation of other physical phenomena like heat and light in terms of mechanical principles. This has further enhanced the importance of mechanics in the evolution of physics.

It should be made clear at the beginning that mechanics is not an explanation of why bodies move. It seeks the simplest possible description of how bodies move. It is true that this description does not content itself with ordinary, common-sense observation expressed in terms of everyday language. Rather, it involves the construction of what has come to be called a physical theory of motion, a logical structure in which certain general but plausible principles are postulated without proof. From these principles physicists can deduce mathematically all the possible motions of The latter can then be compared with actually observed motions, and the extent of the agreement between the theoretically predicted results and those observed is treated as a measure of the success of the theory. In particular if the theory is able to predict a result which has not previously been observed and subsequent test leads to agreement with the prediction, confidence in the success of the theory is considerably enhanced. From this standpoint. theoretical mechanics has proved to be a highly successful physical theory and it is therefore hardly surprising that up to very recent

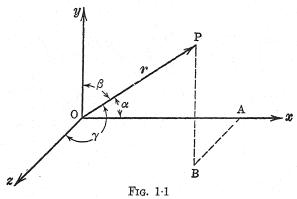
¹ For a discussion of the nature of a physical theory cf. Lindsay and Margenau: "Foundations of Physics" (Wiley, 1936), Chap. I.

times it has been the endeavor of physicists to cast all physical theories into mechanical form, that is, to reduce the description of all natural phenomena in the last analysis to the motions of particles or continuous media of appropriate properties. procedure has been attended with considerable success, though the satisfactory application of it to certain modern problems like the structure of atoms has demanded modifications in the fundamental postulates leading to the generalization of classical mechanics which is called quantum mechanics.

It is the aim of the present volume to set forth the fundamental principles of classical mechanics and to illustrate them with many applications to concrete physical problems of all kinds, not only large-scale observed motions but also phenomena described in terms of the motion of hypothetical particles like molecules, atoms and electrons. In this way, it is hoped that the reader will secure a proper perspective on the value of mechanics as a tool in general physical description.

1.2. Fundamental Definitions. Since mechanics is the science of motion, the question at once arises: what moves, and what are the simplest elements in terms of which its motion may be described? A careful analysis of the first question indicates that a logically consistent answer is really very hard to find; the usual statement that it is matter which moves is not very helpful, since we do not know precisely what matter is, though we are now well acquainted with many of its properties. However, in order to avoid too much of a philosophical discussion, in this book we shall adopt the simple assumption (recognizing, of course, its shortcomings) that the fundamental entity in mechanics is the material particle and that all gross bodies whose changes in position with respect to their surroundings we denominate by the term motion, are aggregates of such particles. At the present stage the simplest and most profitable view which the student can take of such a particle is that it is a geometrical point with the important added nongeometrical property of inertia, the quantitative measure of which we shall later define as its mass. We shall also assume that particles may have certain relations with respect to each other, e.g., gravitation, electrical attraction or repulsion, etc.

We shall therefore begin our study of mechanics with a discussion of the motion of a single particle. Now the elements in terms of which we describe motion are first of all position in space and time. The thoughtful student will undoubtedly wish to examine closely these fundamental concepts. All that we can well say here is that space and time are strictly modes of perception, ways in which we as human beings distinguish objects and events. Each individual thus possesses his own space and time. But for practical purposes there has been devised a kind of public space in which we all agree to conceive the objects of perception to be



placed, and a public time in which events take place. This public space, which is likewise that used in elementary mechanics, is Euclidean and three dimensional. There is one very important fact concerning the position of an object, such as a particle, in this space: we can never know more than its position with respect to other particles, which, of course, may include ourselves as observers. When we speak of the position of a particle, what we really imply is a method of reaching that particle, starting from some given position. Hence position is wholly relative, and implies an arbitrary set of particles or bodies as a reference system. In mechanics this reference system will vary. For many problems it will consist of a point or points on the surface of the earth; for others it will be the sun; while the most general system of this kind is provided by the so-called "fixed" stars. This latter system is known as the primary inertial system.

There are many ways of representing mathematically the posi¹See H. Poincaré, "Foundations of Science" (trans. by G. B. Halsted),
a work which should be in the library of every student of physics. Cf. also

"Foundations of Physics," Chap. II.

4 THE ELEMENTAL CONCEPTS OF MECHANICS

tion of a particle. The two most common ways are (1) by means of rectangular coördinates and (2) by a vector. Consulting Fig. 1·1, the point P may have its position referred to a set of three mutually perpendicular planes intersecting at the point O and in the lines Ox, Oy and Oz (the so-called coördinate axes) respectively, by means of the three perpendicular distances from P to the yz, xz and xy planes respectively. These three distances, usually denoted by x, y and z for convenience, are called the rectangular coördinates of the point P in the reference system O-xyz; and the three values are in themselves sufficient to determine the position of the point.

But it is clear that the point P is also completely specified by its distance r from the point O and the direction of this line with respect to the three coördinate axes, the latter being given by the angles α , β , γ which the line OP makes with these axes. Such a line OP with definite length and in a specified direction is called a vector, for it is a physical quantity which has both magnitude (i.e., as given by its length) and direction (as given by the three angles). We may say that the vector OP completely determines the position of the point P and the particle located there in the given reference system. We shall call OP = r the position vector corresponding to the position P.2 It might at first be thought that the specification of P by means of a vector involves the use of four independent quantities, viz., r and the angles α , β and γ ; whereas the rectangular coördinates are but three in number. However the student will at once recall that α , β , γ are not independent but are connected by the relation

$$\cos^2 \alpha + \cos^2 \beta + \cos^2 \gamma = 1. \tag{1.2-1}$$

As a matter of fact it follows from an examination of the figure that the following connection holds between the rectangular coordinates and the angles, viz.:

$$x = r \cos \alpha$$
, $y = r \cos \beta$, $z = r \cos \gamma$, $(1.2-2)$

² Note the use of bold face type for the designation of vector quantities.

¹ The axes indicated in Fig. 1·1 form what is called a right-handed set in the sense that the positive direction of the z axis is fixed by the direction of advance of a right-handed screw rotated in the xy plane from the positive direction of the x axis to that of the y axis through the angle between them. Interchange of the y and z axes leads to a left-handed set, useful for some purposes. We shall use right-handed axes in this book.

whence, since $x^2 + y^2 + z^2 = r^2$, the identity (1·2-1) follows immediately. The rectangular projection of a vector on a given line is termed the rectangular component (usually more simply component) of the vector along the given line and is equal in magnitude to the product of the magnitude of the vector by the cosine of the angle between the positive directions of the vector and the line. Thus it develops that the rectangular coördinates of P are the components of the vector \mathbf{OP} along the three coördinate axes respectively. (See Sec. 1·3.) We shall find that both methods of locating P are useful, both when P may be considered as lying in a plane and also when three dimensions are essential. The methods of specifying position known as spherical and cylindrical coordinates will be dealt with later.

Specification of position by distance, of course, demands the introduction of a unit of length. In the metric system this is the meter, which is the distance between two marks on a bar of platinum carefully preserved at the International Bureau of Metric Standards, St. Cloud, near Paris. The common unit is the centimeter (cm.) which is 1/100 of a meter. The British unit is the foot, which is 1/3 of the "Standard Yard," the latter being the distance between two marks on a metal bar preserved in London.

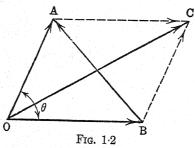
We must next consider the element of time. All physical events, including the motion of particles and bodies, are said to take place The time to which mechanical motions are referred is the public time, which is arbitrarily related to the observed motions of the heavenly bodies, and it has been agreed in mechanics to take as a unit of this time 1/86,400 of a mean solar day. The mean solar day is the time (averaged over one year) between two successive transits of the sun across a given meridian. of time thus defined is called the mean solar second. Actually when in theoretical mechanics it is stated, for example, that the position of a particle is a function of the time, this really means that the position can be most conveniently represented as dependent on an independent variable which can take on an aggregate of values capable of being put into one-to-one correspondence with the real number continuum or, if we like, the points on a line. From this standpoint an interval of time is represented simply as

¹ Astronomers use also the mean sidereal second, which is 1/86,400 of the mean sidereal day or 365.25/366.25 of the mean solar second.

the interval between two real numbers or two points on a line. However, in actual practice these two values are made to correspond with two readings on a clock which registers public time. This is the modern point of view which replaces Newton's definition of "absolute" time as "flowing uniformly." For the latter definition is clearly circular. Nevertheless it is also evident that Newton felt the necessity for the continuity of the abstract time of physics as opposed to the discreteness of the time of perception.

Time like space is relative, and Einstein has shown that it is necessary to modify the common notion regarding the simultaneity of two events. The Einstein modifications are of practical importance only for particles moving with very high velocities (e.g., close to that of light).

The development of physical concepts suitable for physical description, e.g., for use in building a theory like mechanics, affords a problem of great interest in what may be called the foundations of physics. Already much thought has been given to this question. One of the most striking points of view is that of P. W. Bridgman, who stresses what may be called the *operational* standpoint. According to this every physical concept possesses meaning only in so far as there exists a laboratory operation or set



of operations by which a number may be attached to the symbol representing the concept.

1.3. Displacement and Velocity. When a particle changes its position, it is said to undergo a displacement. Since this may be treated as

a change in the position vector of the particle, it is essential that we introduce some considerations on the properties of vectors. Let us consider the two vectors \mathbf{OA} and \mathbf{OB} in the above figure (Fig. 1-2). Through B draw the line parallel to and in the same direction as \mathbf{OA} ; through A draw the line parallel to and in the

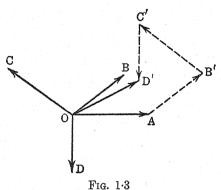
¹ P. W. Bridgman, "The Logic of Modern Physics" (Macmillan Co., New York, 1927). This has much material which will prove very fascinating for the mature student. Cf. also R. B. Lindsay, "Philosophy of Science," 4, 456, 1937, for a critique of operationalism.

same direction as OB. Let these two lines intersect at C. Then the vector OC will be defined as the *sum* or *resultant* of the two vectors OA and OB. Thus we have the *vector* equation

$$OA + OB = OC. (1.3-1)$$

The sum of the two vectors may be looked upon as that diagonal of the parallelogram with **OA** and **OB** as adjacent sides which is included by these two sides. We may also consider the resultant as the third side of the triangle formed by **OB** and **BC**, where **BC** is

equal in magnitude to **OA** and parallel thereto. This point of view is particularly valuable when we desire to find the sum of several vectors: we merely lay off at the end of the first vector a line equal in magnitude to and in the same direction as the second vector. We repeat in turn for all the vectors, and the resultant will be the vector from the origin



to the end of the line last drawn, thus completing a polygon. (Consult the figure — Fig. 1·3 — where OD' is the sum of OA, OB, OC and OD, the four vectors being taken for convenience in the same plane, although this is not necessary for the application of the method.)

Reverting to the simpler case of two vectors, we must be careful to note that eq. $(1\cdot3-1)$ is not an algebraic but a vector equation. The magnitude of the resultant **OC** is given by the cosine law, i.e.,

$$|OC|^2 = |OA|^2 + |OB|^2 + 2|OA| \cdot |OB| \cos \theta, \quad (1.3-2)$$

where the bars indicate absolute value or magnitude.

We are now in a position to discuss the subtraction of two vectors. Consulting Fig. 1·2 once more, we are led to define the vector BA as the *difference* between OA and OB, for it is the vector which when added to OB (by the preceding rule) yields OA. We thus have the vector equation

$$OA - OB = BA. (1.3-3)$$

THE ELEMENTAL CONCEPTS OF MECHANICS 8

The magnitude of the difference vector will again be obtained from the cosine law, i.e.,

$$|AB|^2 = |OA|^2 + |OB|^2 - 2|OA| \cdot |OB| \cos \theta.$$
 (1.3-4)

A very important illustration of the addition of vectors is afforded by the resolution of the position vector r corresponding to a given point (Sec. 1.2) into its components along three mutually perpendicular lines, i.e., the rectangular components x, y, z. If we consult again Fig. 1.1, the law of vector addition gives us at once the vector equation

$$OP = OA + AB + BP, (1.3-5)$$

where OP = r, while |OA| = x, |BP| = y and |AB| = z. If now we denote by i, j, k vectors of unit magnitude along the x, y and zdirections respectively, we may write

$$OA = ix$$
, $BP = jy$, $AB = kz$,

whence (1.3-5) becomes

$$\mathbf{r} = \mathbf{i}x + \mathbf{j}y + \mathbf{k}z. \tag{1.3-6}$$

This mode of representing vectors in terms of their rectangular

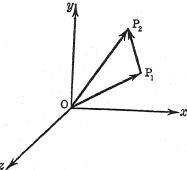


Fig. 1.4

components is of considerable utility.

It is now in order to discuss the displacement of a particle. Consulting Fig. 1.4 we note that the particle is conceived at first to be at the point P_1 and at a short time interval Δt later to be at the point P_2 . If the displacement has taken place along the line P_1P_2 it is best represented by the vector element P_1P_2 which is equal to

the difference between the vectors giving the respective positions P_2 and P_1 , i.e., we have

$$P_1P_2 = OP_2 - OP_1 = r_2 - r_1,$$
 (1.3-7)

where r_1 and r_2 are the position vectors corresponding to the positions P_1 and P_2 respectively. We shall then write

$$\mathbf{P}_{1}\mathbf{P}_{2} = \Delta \mathbf{r} \tag{1.3-8}$$

as a simplified expression for the displacement. We shall now define $\frac{\Delta \mathbf{r}}{\Delta t}$ as the average velocity of displacement from P_1 to P_2 .

It is the average time rate of change of position of the particle from its original position P_1 to its final position P_2 . It will be noted that we have here confined our attention to an elementary displacement. But we can easily apply the same method to any finite displacement along any path whatever by considering the path to be made up of a great number of such elementary displacements.

The instantaneous velocity of the particle at the position P_1 is then defined as follows:

$$\lim_{\Delta t = 0} \frac{\Delta \mathbf{r}}{\Delta t} = \frac{d\mathbf{r}}{dt} = \mathbf{v},\tag{1-3-9}$$

which, of course, assumes that the derivative actually exists. This will be the case in all the problems in mechanics which we shall study. It is clear that velocity like displacement is a vector quantity; its magnitude is often referred to as the speed and will be denoted simply by v. In the metric system it will commonly be expressed in cm/sec, while in the British system the unit is the foot/sec.

It is often convenient to resolve a velocity into components along the x, y and z axes in a coördinate system. We shall denote these three components by v_x , v_y and v_z respectively. Let Δx , Δy , Δz be the components of Δr along the coördinate axes, i.e.,

$$\Delta x = |\Delta r| \cos \xi,$$

$$\Delta y = |\Delta r| \cos \eta,$$

$$\Delta z = |\Delta r| \cos \zeta,$$
(1.3-10)

where ξ , η and ζ are the angles Δr makes with the three axes respectively. We then have

$$\lim_{\Delta t = 0} \frac{\Delta x}{\Delta t} = \frac{dx}{dt} = \dot{x} = v_x,$$

$$\lim_{\Delta t = 0} \frac{\Delta y}{\Delta t} = \frac{dy}{dt} = \dot{y} = v_y,$$

$$\lim_{\Delta t = 0} \frac{\Delta z}{\Delta t} = \frac{dz}{dt} = \dot{z} = v_z.$$
(1.3–11)

10 THE ELEMENTAL CONCEPTS OF MECHANICS

Attention should be called to the use of the dot to indicate differentiation with respect to the time. Two dots will indicate the second derivative, etc. This notation will be used throughout the book because of its simplicity. It will be noted that v_x , v_y and v_z are no longer vectors, for they are the magnitudes of the components of the vector velocity along the coördinate axes. Since

 $\cos^2 \xi + \cos^2 \eta + \cos^2 \zeta = 1,$

we have

$$|\mathbf{v}|^2 = v^2 = v_x^2 + v_y^2 + v_z^2.$$
 (1.3-12)

The representation of a velocity vector in terms of its components is termed the *resolution* of the vector. From another point of view we can consider the vector as compounded of three mutually perpendicular vectors of magnitudes equal to the components v_z , v_y and v_z respectively, by the general method of summation of vectors discussed in the first part of this section. For velocity vectors follow the same rules for addition and subtraction as position and displacement vectors. Thus corresponding to eq. (1.3-6) we shall have here

$$\mathbf{v} = \dot{\mathbf{r}} = \mathbf{i}v_x + \mathbf{j}v_y + \mathbf{k}v_z. \tag{1.3-13}$$

In fact the student should realize that the vector formulas we have been using in this section are perfectly general in form, i.e., hold for vectors in general. Anyone interested in vector analysis as a branch of mathematics may turn with profit to a text like that of J. G. Coffin, *Vector Analysis* (Wiley, New York, 2nd ed., 1911).

There is, indeed, one important point to note concerning the composition of vector velocities by the parallelogram or polygon rule, namely, that in using the latter in this case we are really making an assumption of a very fundamental nature, though one not often emphasized in elementary texts. The assumption is that the velocities are mutually independent, i.e., the fact that a particle has the velocity \mathbf{v}_1 does not influence the imposition of a velocity \mathbf{v}_2 at the same time. This mutual independence is not universal in physics, but where the assumption is justified, as in the present instance, it introduces marked simplification. The

¹ Cf. also H. B. Phillips, "Vector Analysis" (Wiley, New York, 1933).

² The student may recall from elementary electrostatics that the mutual influence of two charged conductors is *not* independent of the presence of a third conductor.

hypothesis is usually referred to as the principle of superposition. We shall meet with other illustrations.

1.4. Linear and Angular Motion. The notions of displacement and velocity developed in the preceding sections are quite general. But there is a particular method of measuring displacement and velocity which is often of such great use that special mention must be made of it here preparatory to the more thorough discussion of Chapters II and III. Let us consider the special

case of the motion of a particle in a circle. Examining the figure (Fig. 1.5) we note that the motion of the particle from P to Q may be described either by the distance Δs traveled along the circumference or by the angle $\Delta \theta$ through which the line joining the position of the particle to the center of the circle is displaced. In physics the common unit of such angular displacement is the radian, which is the angle at the center of a

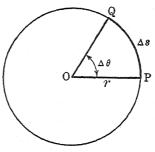


Fig. 1.5

circle subtended by an arc equal in length to the radius. It follows immediately that the one displacement may be expressed in terms of the other by the relation

$$\Delta s = r \, \Delta \theta. \tag{1.4-1}$$

If the displacement from P to Q takes place in an interval of time Δt , the average angular velocity during the displacement may be defined as $\Delta \theta/\Delta t$ and the instantaneous angular velocity at P will then be

$$\omega = \lim_{t \to 0} \frac{\Delta \theta}{\Delta t} = \frac{d\theta}{dt} = \dot{\theta}, \qquad (1.4-2)$$

expressed in radians per second. From (1·4-1) it follows that the corresponding instantaneous velocity along the arc of the circle,

viz.,
$$v = \frac{ds}{dt} = \dot{s}$$
 is related to ω by the relation

$$v = r\omega, \tag{1.4-3}$$

where v will be expressed in cm/sec if r is in cm and ω in radians

12 THE ELEMENTAL CONCEPTS OF MECHANICS

per sec. If the angular velocity is expressed, as is sometimes done, in revolutions per sec, say n in number, we have

$$\omega = 2\pi n, \qquad (1.4-4)$$

and therefore

$$v = 2\pi n r, \qquad (1.4-5)$$

expressions which the student will have frequent occasion to use. It may be noted, of course, that the above relationships are strictly scalar in nature. Yet velocity along a curve is a vector, as we have seen. Is then angular velocity representable by a vector? It is, and the representation is usually made by a line drawn from the center of the circle O perpendicular to the plane of the motion directed in such a way that as one looks along the vector from O the sense of revolution is clockwise. Alternatively one may say that the direction of the vector is the direction of advance of a right-handed screw the sense of whose revolution is the same as that of the given motion. This provides a rule which it is particularly simple to remember. The length of the line is proportional to the magnitude of the angular velocity. As a matter of fact, finite angular displacements are not compoundable in the same way as linear displacement or velocity vectors, and hence we shall not consider them as vectors in this book. (See Problem 11 at the end of the chapter.) However, finite angular velocities are compoundable in the usual sense and as vectors turn out to be of importance in the motion of rigid bodies.

1.5. Acceleration. The definition of instantaneous velocity given in Sec. 1.3 implies that the velocity of a particle may change with the time. The time rate of change of velocity is called the acceleration of the particle and is one of the fundamental quantities of mechanics. Thus, if in the time element Δt the linear velocity of a particle has changed from \mathbf{v}_1 to \mathbf{v}_2 , we have

$$\mathbf{a}_{av} = \frac{\mathbf{v}_2 - \mathbf{v}_1}{\Delta t} = \frac{\Delta \mathbf{v}}{\Delta t}, \qquad (1.5-1)$$

and similarly

$$\mathbf{a}_{\text{inst}} = \lim_{\Delta t = 0} \frac{\Delta \mathbf{v}}{\Delta t} = \frac{d\mathbf{v}}{dt} = \dot{\mathbf{v}}.$$
 (1.5-2)

Since \mathbf{v} , the instantaneous velocity, has been defined in Sec. 1·3 as $\frac{d\mathbf{r}}{dt}$, we can write

$$\mathbf{a} = \frac{d}{dt} \left(\frac{d\mathbf{r}}{dt} \right) = \frac{d^2\mathbf{r}}{dt^2} = \ddot{\mathbf{r}}, \tag{1.5-3}$$

thus defining a as the second derivative of the position vector r with respect to the time. It is clear that acceleration is a vector since

the difference between two velocities is a vector. It may be observed that for this reason, even though the magnitude of the velocity of a particle may not change, it may nevertheless experience an acceleration due to change in the direction of its velocity. The classical illustration of this is to be found in uniform motion in a circle. We can treat this in simple fashion by an investigation of the adjacent

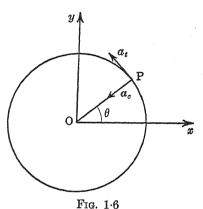


figure (Fig. 1.6). Let the particle move with constant angular velocity ω radians/second in the circle of radius r with center at O. The position vector of point P on the circular path is (cf. 1.3–6)

$$\mathbf{r} = \mathbf{i}x + \mathbf{j}y = r(\mathbf{i}\cos\theta + \mathbf{j}\sin\theta).$$
 (1.5-4)

Hence by differentiation with respect to the time, remembering that r is constant, the instantaneous velocity becomes

$$\mathbf{v} = \dot{\mathbf{r}} = r(-\mathbf{i}\sin\theta + \mathbf{j}\cos\theta)\dot{\theta}, \tag{1.5-5}$$

leading to the speed

$$v = |\mathbf{v}| = r\dot{\theta} = r\omega, \tag{1.5-6}$$

in agreement with (1.4-3). A second time differentiation yields $\mathbf{a} = \dot{\mathbf{v}} = \ddot{\mathbf{r}} = r(-\mathbf{i}\cos\theta - \mathbf{j}\sin\theta)\dot{\theta}^2 + r(-\mathbf{i}\sin\theta + \mathbf{j}\cos\theta)\ddot{\theta}.$ (1.5-7)

This is more conveniently written

$$\mathbf{a} = r\mathbf{i}(-\cos\theta \cdot \dot{\theta}^2 - \sin\theta \cdot \ddot{\theta}) + r\mathbf{j}(-\sin\theta \cdot \dot{\theta}^2 + \cos\theta \cdot \ddot{\theta}).$$
(1.5-8)

14 THE ELEMENTAL CONCEPTS OF MECHANICS

In the usual way the magnitude of the acceleration becomes

$$a = \sqrt{a_x^2 + a_y^2} = r\sqrt{\dot{\theta}^4 + \ddot{\theta}^2}.$$
 (1.5-9)

To give a physical interpretation to (1.5-9) let us assume first that $\ddot{\theta} = 0$. This means that $\dot{\omega} = 0$, i.e., there is no change in the magnitude of the angular velocity with the time. We shall say that there is no angular acceleration, putting

$$\ddot{\theta} = \alpha, \tag{1.5-10}$$

for convenience. Then

$$a = r\dot{\theta}^2 = r\omega^2. \tag{1.5-11}$$

Thus even when the angular speed ω is constant the particle still has an acceleration whose magnitude is given by (1·5–11). It remains to be ascertained what the direction of this acceleration is. From (1·5–8) (we are still considering $\ddot{\theta}=0$) we have

$$a_x = -r\dot{\theta}^2\cos\theta, \quad a_y = -r\dot{\theta}^2\sin\theta.$$
 (1.5-12)

But these are precisely the components of a vector of magnitude $r\theta^2$ directed along the radius r toward the center of the circle O. Hence the acceleration a is called the *centripetal* acceleration, with magnitude denoted by a_c .

It then develops that a particle moving in a circle is subject to two kinds of acceleration. Whether the angular velocity is constant or not, there always exists the centripetal acceleration with magnitude a_c . If in addition there is an angular acceleration α , there exists another acceleration of magnitude $r\alpha$. The direction of this can be seen from an examination of (1.5–8), since the components are, respectively,

$$-r\sin\theta\cdot\ddot{\theta}$$
, $r\cos\theta\cdot\ddot{\theta}$

These are the components of a vector of magnitude $r\alpha$ directed tangentially to the circle at P in the direction in which P is moving (i.e., increasing θ). We therefore write

$$a_t = r\alpha, (1.5-13)$$

and refer to this acceleration as the *tangential* acceleration of P. It is clear from (1.5-9) that if a is the magnitude of the resultant acceleration

$$a = \sqrt{a_c^2 + a_t^2}. (1.5-14)$$

1.6. Concept of Inertia and Definition of Mass. So far we have concerned ourselves solely with kinematics, the geometry of motion. This treats particles essentially as geometrical points and neglects any differences that may exist between them. Actually two particles that look alike may move quite differently even when subjected to what appear to be identical conditions. To explain this the concept of inertia has been invented with mass as its quantitative measure. This is inextricably related with the mutual motion of two or more particles as will be made clear in the definition we shall now examine.

Let us suppose that we have two particles A and B which are completely isolated from all others in the universe. Under their mutual action they will suffer accelerations which with reference to the primary inertial system we may denote by a_{AB} and a_{BA} respectively. Now all experiments which have ever been performed on actual bodies lead us to assume that in such a case as we are discussing, these accelerations are opposite in direction and that their ratio is a scalar constant which is independent of the relative positions of the particles, of their velocities (save when the latter approach in value the velocity of light) and the time. We shall extend this assumption by supposing that this ratio is also independent of the presence of other particles, where, of course, it is to be clearly understood (as the above notation indicates) that the accelerations in question are due solely to mutual action. The presence of other particles will indeed produce accelerations in A and B, but our assumption is that the ratio of the acceleration of A due to B to that of B due to A is not influenced by such. We thus have, no matter what the magnitudes of the individual accelerations may be,

$$\frac{\mathbf{a}_{AB}}{-\mathbf{a}_{BA}} = K_{BA}. \tag{1.6-1}$$

The constancy of this ratio may be looked upon as one of the most fundamental laws of mechanics. We are now free to give some interpretation to the magnitude of the constant K_{BA} . Let us take any other particle C. For the mutual actions of C and A, and C

¹ Cf. "Foundations of Physics," p. 91 ff. Also H. Poincaré, "Science and Hypothesis" (English translation in "Foundations of Science"), p. 98 ff. Also Ernst Mach, "The Science of Mechanics" (English translation by Open Court Publishing Co., Chicago), p. 216 ff.

16 THE ELEMENTAL CONCEPTS OF MECHANICS

and B respectively we shall then have analogously to (1.6-1)

$$\frac{\mathbf{a}_{AC}}{-\mathbf{a}_{CA}} = K_{CA}, \quad \frac{\mathbf{a}_{BC}}{-\mathbf{a}_{CB}} = K_{CB}. \tag{1.6-2}$$

Now experiment indicates that the following relation holds between the three constants K_{BA} , K_{CA} and K_{CB} , namely,

$$K_{CB} = \frac{K_{CA}}{K_{BA}}$$
 (1.6-3)

The second of the relations (1.6-2) then becomes

$$K_{BA}\mathbf{a}_{BC} = -K_{CA}\mathbf{a}_{CB}. \tag{1.6-4}$$

The relation (1.6-4) indicates that we are able by choosing A as a standard particle to associate with every other particle a constant which does not depend at all on the particle with which the one in question interacts. Its importance will be emphasized if we give to it or some simple function of it a name. This may, of course, be done in a variety of ways, but we shall prefer to name the constant K_{BA} the mass of B relative to A, or more simply the mass of B, it being understood that A is to be taken as a reference particle. Thus we write $K_{BA} = m_B$, etc., so that (1.6-4) becomes

$$m_B \mathbf{a}_{BC} = -m_C \mathbf{a}_{CB}. \tag{1.6-5}$$

We make further the important assumption also suggested by experiment that if a particle B is influenced by several particles C, D, E, \ldots its resultant acceleration will be the vector sum of the accelerations due to the particles independently, viz.

$$\mathbf{a}_{B,CDE} = \mathbf{a}_{BC} + \mathbf{a}_{BD} + \mathbf{a}_{BE} + \cdots.$$

This is another illustration of the principle of superposition already mentioned in Sec. 1.3. We shall see in the next section that the foregoing definition of mass implies the important additive property that the mass of the particle formed by joining any two is equal to the sum of the individual masses. Indeed this may be looked upon as one reason for choosing the present definition, which is, of course, really arbitrary in nature.¹

In order to measure mass by means of eq. (1.6-5) it is necessary to establish a *unit* by arbitrarily assigning unit mass to a certain volume of some substance. In the metric system the unit of

¹ Thus we might have named $1/K_{BA}$ the mass of B with equal consistency.

mass is the gram, which is defined to be the one-thousandth part of the mass of a lump of platinum carefully preserved in Paris — the so-called standard kilogram. It was originally intended that the standard kilogram should be the mass of 1000 cubic centimeters of water at its maximum density. Actually there is a slight discrepancy so that the volume of one kilogram of water at maximum density is 1000.027 ± 0.001 cubic centimeters. The British unit is the standard pound, whose relation to the kilogram is expressed in the statement that one kilogram is approximately equivalent to 2.205 pounds.

The definition of mass presented in this section is general in nature and the student will find it desirable to think out for himself certain special cases. For example, two particles may be conceived to be connected by a perfectly elastic spring which itself has no mass. When the spring is stretched and released the relation (1·6-1) will be verified irrespective of the extent of stretching. Moreover with other particles and springs, eq. (1·6-3) will also be verified. Other modes of interaction will occur to the mind, such as the mutual actions of gravitating and electrified particles. Our fundamental assumption is that in every case of this sort the relations (1·6-1) and (1·6-3) will be found to be satisfied.

The student is particularly urged to avoid such statements as "mass is the amount of matter in a body." It is becoming increasingly clear that the expression "amount of matter" without further qualification is physically meaningless. The only logical definition of mass is that which enables us to establish a method of comparing masses. This is precisely the object of the definition of this section. Of course, it must be at once admitted that in practice masses are not usually compared in the ideal way indicated. For this purpose it is much more convenient to employ a balance, which utilizes the gravitational attraction of the earth for all bodies. The connection between this type of measurement and the logical definition of mass will be discussed later.

1.7. Force and the Laws of Motion. We are now in a position to introduce formally the fundamental laws of motion resulting from the experimental researches of Galileo and the more mature

¹ See R. T. Birge: "Probable Values of the General Physical Constants" in *Reviews of Modern Physics*, Vol. 1, p. 11, 1929.

considerations of Huygens and Newton, the three greatest physicists of the 17th century. As a matter of fact we have already progressed far toward an understanding of these laws in the preceding section. For let us go back to eqs. (1.6-1) and (1.6-3) with the resulting equation (1.6-4). We were there led to a definition of mass, wherein mass and acceleration for two interacting particles B and C are connected by the relation

$$m_B \mathbf{a}_{BC} = -m_C \mathbf{a}_{CB}$$
.

We now wish to attach further significance to this relation. It states that in the interaction of the two particles there is something which is the same in magnitude for both particles. This is, of course, the product of mass (as we have defined it) and acceleration. We shall now further assume that its value in any given case depends on the relative positions of the particles, though it may conceivably sometimes be a function of the relative velocities of the particles as well as of the time. It is convenient to express this dependence by introducing a new vector function $F_{BC}(x, y, z)$, where x, y, z are the coördinates of the one particle relative to the other, and placing

$$m_B \mathbf{a}_{BC} = \mathbf{F}_{BC}(x, y, z). \tag{1.7-1}$$

We now define the function F_{BC} as the force acting on the particle B due to the influence of the particle C. For C we have similarly

$$m_C \mathbf{a}_{CB} = \mathbf{F}_{CB}, \tag{1.7-2}$$

and the relation (1.6-5) shows us that

$$\mathbf{F}_{BC} = -\mathbf{F}_{CB}, \tag{1.7-3}$$

which means that the force on B due to C is equal and opposite in direction to the force on C due to B. It is to be noted that the force is a vector and is moreover directed along the acceleration. Using the derivative notation we may write $(1\cdot7-1)$ in general form

$$\mathbf{F} = m\dot{\mathbf{v}} = m\ddot{\mathbf{r}} \tag{1.7-4}$$

if $\mathbf{v} = \dot{\mathbf{r}}$, as in Sec. 1.3. This may be called the "equation of motion" of the particle, which is thus obtained by placing the

¹ For simplicity of notation we shall express this here as a function of position only, and leave more complicated cases for later discussion.

UNITS 19

force equal to the product of mass and acceleration, i.e., the so-called *kinetic reaction*. It is fundamental in the solution of all problems in motion.

The definition of force given above is probably the only logically consistent one available in mechanics. For the notions of "push" and "pull" are too purely anthropomorphic for any exact use, though we shall probably continue to employ them for illustrative purposes. In a later chapter we shall see how static force is related to the present definition. The only kind of motion in which we conceive no force to act is clearly that in which the acceleration is zero everywhere and at all times. This is either rest or uniform motion (i.e., with constant velocity) in a straight line. With this addition to our previous discussion we have developed the essential content of Newton's three laws of motion, stated by him in the following form:

- (I) Every body continues in its state of rest or of uniform motion in a straight line, except in so far as it is compelled by forces to change that state.
- (II) Change of motion is proportional to the force and takes place in the direction of the straight line in which the force acts.
- (III) To every action there is always an equal and contrary reaction; or, the mutual actions of any two bodies are always equal and oppositely directed along the same straight line.

In comparing the first two formal statements with the discussion of the present section we must bear in mind that as a measure of the "motion" of a body, Newton used the product of mass and velocity, which in present day usage is termed the momentum, though Newton referred to it as "quantity of motion." Hence Newton's statement of the second law would logically be written in mathematical form

$$\mathbf{F} = \frac{d(m\mathbf{v})}{dt} \cdot \tag{1.7-5}$$

However, if the mass remains constant in time, eq. (1.7-5) is equivalent to eq. (1.7-4). We shall in general prefer (1.7-4).

1.8. Units. The units of mass in the metric and British systems have been stated to be the gram and the pound respectively. The

¹ Cf. "Foundations of Physics," p. 85 ff.

unit of force may be at once defined from eq. (1.7-1). Thus the force which corresponds to one gram of mass moving with an acceleration of one centimeter per second per second is the unit force in the metric or centimeter-gram-second (c.g.s.) system, and is called the dyne. Similarly the force associated with one pound of mass moving with an acceleration of one foot per second per second is called one poundal. Engineers rarely use this unit, preferring to define their unit of force in terms of weight. Thus they call the force with which the earth accelerates a body of mass one pound. a force of one pound. This is sometimes called the gravitational unit of force. Since it is experimentally true that every body in the neighborhood of the earth's surface experiences an approximately constant downward acceleration of "g" (about 32.2 feet/sec² or 980 cm/sec², though variable from place to place¹ within certain narrow limits) it follows that the pound of force or weight is equivalent to "g" poundals. The poundal has the advantage of being an absolute unit, while the pound of force (written in this text hereafter as pound weight and abbreviated to lb wt) varies in magnitude from place to place on the earth's surface, being greatest at the poles and least at the equator. While the student of mechanics is advised to use mainly the absolute units, he should be prepared to transfer his results into the engineering units, for the latter are the most widely used in daily life, at least in English-speaking countries. In the metric system there is an analogous usage, the gram of force or weight being defined as that force with which the earth accelerates a mass of one gram. The gram weight is then "g" dynes (g in metric units here, of course).

The reader will already have noted that the introduction of units is a matter of convention and bound to be arbitrary. For pure science usage the metric c.g.s. (centimeter-gram-second) system has been pretty generally adopted throughout the world. Engineers usually employ some kind of gravitational unit. Much pressure is being exerted on both engineers and scientists to adopt a uniform system. The one commonly urged for this purpose is the so-called m.k.s. (meter-kilogram-second) system. In this the meter is the unit of length, the kilogram the unit of mass, while the second remains the unit of time. The unit of force is the newton

 $^{^1\,\}rm The$ value of "g" at the poles is 983.21 cm/sec² while at the equator it is 977.99 cm/sec².

or one kilogram meter per second per second. It is clearly equal to 10⁵ dynes and hence, so far as magnitude is concerned, a more practical unit than the dyne. Indeed, the m.k.s. system is sometimes called the absolute practical system. It has some advantages in connection with the relation between mechanical and electrical units. In the meantime the reader is advised against a slavish adherence to any particular system. The main thing to keep in mind is the necessity for consistency: in the solution of numerical problems in mechanics, adopt at the start the system of units which seems most convenient and stick to it throughout the analysis.

1.9. The Third Law of Motion. Combination of Masses. first two laws of motion were completely implied in the truly remarkable investigations of Galileo. The third law was the peculiar contribution of Newton himself. We note that it is already present by implication in eq. (1.6-5), and following upon the definition of force the complete mathematical statement of the law was given in eq. (1.7-3). It thus appears as an even more fundamental law of motion than the first two. To state it again in words, it means that whenever one body is accelerated there must be another body accelerated in the opposite direction. Thus accelerations never occur singly but always in pairs. Stated in terms of force we may say that if a force acts on a given body, this body exerts an equal and oppositely directed force on some other body. The significance of the law is perhaps most easily understood with reference to the push and pull forces with which we feel ourselves so familiar. When a book lies on the table we say it presses down on the table with a certain force: the law states then that the table pushes up with precisely the same force. The locomotive pulls forward on a train of cars and the cars pull back on the locomotive with precisely the same force, irrespective of the actual state of motion of the whole train. Newton called the two aspects of the force action and reaction, whence the formal statement of the law given above.

The recoil of a gun is another familiar illustration of the principle. That the latter is sometimes a matter of confusion is evident from the well known student paradox: "if action and reaction are equal and opposite, how can bodies ever move?" The well-worn illustration of the horse pulling the wagon will suffice to clear up this point: the horse pulls forward on the wagon and

the wagon pulls back on the horse with the same force, yet the wagon moves! The answer is that the acceleration of the wagon is a function only of the forces acting on the wagon, which in this instance are the forward pull of the horse and the backward pull of the ground. The force which the wagon itself exerts has nothing to do with its motion — this force concerns the motions of other bodies only, for example, the horse in our illustration.

In the previous section it was pointed out that the definition of mass there presented implies the additive property that the mass of two particles joined together is equal to the sum of the individual masses. The truth of this statement is now seen to follow from the third law of motion.

Consider the three mass particles B, C and D and suppose that B and C have approached so closely to each other that they may effectively be considered as forming a single particle under the influence of D. The equation (1.6-5) then becomes

$$m_{(C+B)}\mathbf{a}_{(C+B),D} = -m_D\mathbf{a}_{D,(C+B)},$$
 (1.9-1)

where $m_{(C+B)}$ denotes the joint mass of C and B considered as a single particle, and $\mathbf{a}_{(C+B),D}$ is the acceleration of the joint particle due to D. The left-hand side is equal by definition to the force which D exerts on the joint particle C+B. Call this $\mathbf{F}_{(C+B),D}$. By the use of the principle of superposition this may be considered as the resultant of the forces which D exerts on C and B separately. Thus

$$\mathbf{F}_{(C+B),D} = \mathbf{F}_{CD} + \mathbf{F}_{BD}.$$
 (1.9-2)

Now no matter what the mutual orientation of B, C, D is, the total force on C is the vector sum of the force due to D and that due to B and this is equal to the mass of C times the acceleration of C, since we have already assumed in the preceding section that the resultant acceleration is the vector sum of the individual accelerations. There is a similar equation for B. We may write them both as follows:

$$\mathbf{F}_{CD} + \mathbf{F}_{CB} = m_C \mathbf{a}_{C,BD},$$
 (1.9-3)

$$\mathbf{F}_{BD} + \mathbf{F}_{BC} = m_B \mathbf{a}_{B,CD}, \qquad (1.9-4)$$

where \mathbf{F}_{CB} and \mathbf{F}_{BC} are respectively the force on C due to B and that on B due to C. From the third law of motion, we have, of

course, no matter of what character the action may be,

$$F_{CB} + F_{BC} = 0.$$
 (1.9-5)

Hence if we add (1.9-3) and (1.9-4) having regard for (1.9-5) we get

$$\mathbf{F}_{CD} + \mathbf{F}_{BD} = m_C \mathbf{a}_{C,BD} + m_B \mathbf{a}_{B,CD},$$
 (1.9-6)

where $\mathbf{a}_{C,BD}$ and $\mathbf{a}_{B,CD}$ are respectively the accelerations of C due to B and D, and of B due to C and D. For any arbitrary orientation of B, C and D these accelerations may be quite different. However, since we are going to assume that B and C are to form the equivalent of a single particle, we must suppose that by the time we are considering them they move in such a way that their distance apart always remains within certain fixed limits, the upper of which is very small compared with the distance of either from D. To specify exactly their mutual motion under this condition would demand a precise assumption with respect to the nature of the forces between them and this, of course, we desire to avoid since we wish our result to be as general as possible. However, if we finally neglect small quantities of the second order, which is equivalent to assuming that the distance between B and C remains invariable, we may write

$$a_{C,BD} = a_{B,CD} = a_{(C+B),D}.$$
 (1.9-7)

Consequently eq. (1.9-6) becomes

$$\mathbf{F}_{CD} + \mathbf{F}_{BD} = (m_C + m_B) \mathbf{a}_{(C+B),D}.$$
 (1.9-8)

It now follows from (1.9-1) and (1.9-2) that

$$m_{(C+B)}\mathbf{a}_{(C+B),D} = (m_C + m_B)\mathbf{a}_{(C+B),D},$$

whence

$$m_{(C+B)} = m_C + m_B, (1.9-9)$$

and the additive property is thus confirmed.

The student may well inquire what attitude he should adopt toward the laws of motion, i.e., are they truly experimental laws or are they merely postulates — a part of the fundamental theory of mechanics? Careful consideration of this question by such men as K. Pearson, H. Poincaré, Ernst Mach, H. Hertz, and others indicates that the latter attitude is the more strictly logical one, for while the laws are of course suggested by experiment, no experi-

the wagon pulls back on the horse with the same force, yet the wagon moves! The answer is that the acceleration of the wagon is a function only of the forces acting on the wagon, which in this instance are the *forward* pull of the horse and the *backward* pull of the ground. The force which the wagon *itself* exerts has nothing to do with *its* motion — this force concerns the motions of other bodies only, for example, the horse in our illustration.

In the previous section it was pointed out that the definition of mass there presented implies the additive property that the mass of two particles joined together is equal to the sum of the individual masses. The truth of this statement is now seen to follow from the third law of motion.

Consider the three mass particles B, C and D and suppose that B and C have approached so closely to each other that they may effectively be considered as forming a single particle under the influence of D. The equation $(1\cdot6-5)$ then becomes

$$m_{(C+B)}\mathbf{a}_{(C+B),D} = -m_D\mathbf{a}_{D,(C+B)},$$
 (1.9-1)

where $m_{(C+B)}$ denotes the joint mass of C and B considered as a single particle, and $\mathbf{a}_{(C+B),D}$ is the acceleration of the joint particle due to D. The left-hand side is equal by definition to the force which D exerts on the joint particle C+B. Call this $\mathbf{F}_{(C+B),D}$. By the use of the principle of superposition this may be considered as the resultant of the forces which D exerts on C and B separately. Thus

$$\mathbf{F}_{(C+B),D} = \mathbf{F}_{CD} + \mathbf{F}_{BD}.$$
 (1.9-2)

Now no matter what the mutual orientation of B, C, D is, the total force on C is the vector sum of the force due to D and that due to B and this is equal to the mass of C times the acceleration of C, since we have already assumed in the preceding section that the resultant acceleration is the vector sum of the individual accelerations. There is a similar equation for B. We may write themboth as follows:

$$\mathbf{F}_{CD} + \mathbf{F}_{CB} = m_C \mathbf{a}_{C,BD},$$
 (1.9-3)

$$F_{BD} + F_{BC} = m_B a_{B,CD},$$
 (1.9-4)

where \mathbf{F}_{CB} and \mathbf{F}_{BC} are respectively the force on C due to B and that on B due to C. From the third law of motion, we have, of

course, no matter of what character the action may be,

$$\mathbf{F}_{CB} + \mathbf{F}_{BC} = 0.$$
 (1.9-5)

Hence if we add (1.9-3) and (1.9-4) having regard for (1.9-5) we get

$$\mathbf{F}_{CD} + \mathbf{F}_{BD} = m_C \mathbf{a}_{C,BD} + m_B \mathbf{a}_{B,CD},$$
 (1.9-6)

where $\mathbf{a}_{C,BD}$ and $\mathbf{a}_{B,CD}$ are respectively the accelerations of C due to B and D, and of B due to C and D. For any arbitrary orientation of B, C and D these accelerations may be quite different. However, since we are going to assume that B and C are to form the equivalent of a single particle, we must suppose that by the time we are considering them they move in such a way that their distance apart always remains within certain fixed limits, the upper of which is very small compared with the distance of either from D. To specify exactly their mutual motion under this condition would demand a precise assumption with respect to the nature of the forces between them and this, of course, we desire to avoid since we wish our result to be as general as possible. However, if we finally neglect small quantities of the second order, which is equivalent to assuming that the distance between B and C remains invariable, we may write

$$\mathbf{a}_{C,BD} = \mathbf{a}_{B,CD} = \mathbf{a}_{(C+B),D}.$$
 (1.9-7)

Consequently eq. (1.9-6) becomes

$$\mathbf{F}_{CD} + \mathbf{F}_{BD} = (m_C + m_B)\mathbf{a}_{(C+B),D}.$$
 (1.9-8)

It now follows from (1.9-1) and (1.9-2) that

$$m_{(C+B)}\mathbf{a}_{(C+B),D} = (m_C + m_B)\mathbf{a}_{(C+B),D},$$

whence

$$m_{(C+B)} = m_C + m_B, (1.9-9)$$

and the additive property is thus confirmed.

The student may well inquire what attitude he should adopt toward the laws of motion, i.e., are they truly experimental laws or are they merely postulates — a part of the fundamental theory of mechanics? Careful consideration of this question by such men as K. Pearson, H. Poincaré, Ernst Mach, H. Hertz, and others indicates that the latter attitude is the more strictly logical one, for while the laws are of course suggested by experiment, no experi-

ment can ever be devised to test them fully and unambiguously. They may be taken to represent our idea as to the most general and simple basis, compatible with experiment, for the development of mechanics. Beginning with them as assumptions we shall make various deductions which may then be checked by the results of the most careful experiments which can be carried out. We shall see in the course of our work that beside the three laws of Newton, other more compact formulations of the principles of mechanics are possible and very valuable for certain purposes.

1.10. Impulse and Kinetic Energy. Work and Power. In addition to mass and force there are certain other mechanical concepts so important that it seems desirable to introduce them here, even though we shall postpone their detailed application to Chapter IV.

What is the cumulative effect of a force acting on a particle as time passes? The natural measure of this is the time integral of the force over the interval in question. This is known as the *impulse* of the force in this interval. A simple but significant theorem can be proved about this quantity. Thus if we employ the equation of motion (1.7-5) we can write

$$\int_{t_0}^{t_1} \mathbf{F} dt = m \int_{t_0}^{t_1} \frac{d\mathbf{v}}{dt} dt = m \int_{\mathbf{v}_0}^{\mathbf{v}_1} d\mathbf{v}, \qquad (1.10-1)$$

where v_0 and v_1 are the velocities of the particle at the instants t_0 and t_1 respectively. Integration yields

$$\int_{t_0}^{t_1} \mathbf{F} \, dt = m \mathbf{v}_1 - m \mathbf{v}_0. \tag{1.10-2}$$

It thus appears that the impulse of the resultant force is the *change* in *momentum* of the particle during the interval in which the force acts. This is known as the impulse-momentum theorem. It is of particular interest for forces which are of short duration.

For example, though the force may be very great (i.e., mathematically approach infinity) the time integral $(1\cdot10-2)$ may nevertheless be finite if the time interval is small enough. Thus we may have for $\mathbf{F} \doteq \infty$ and $\tau \doteq 0$ for a particle of mass m

$$m(\mathbf{v}_2 - \mathbf{v}_1) = \int_0^{\tau} \mathbf{F} dt = \mathbf{J},$$
 (1.10-3)

where v_1 and v_2 are the velocities of the particle at the beginning and end of the time interval τ and J is a finite vector quantity. A very interesting though curious feature of this problem is seen in the following. Suppose the velocity of the particle remains always finite and let v_{max} be its largest value during the interval. Then the actual distance traveled during the duration of application of the force is certainly less than τv_{max} . But as $\tau \doteq 0$ this approaches zero, and hence we conclude that the effect of this infinite force for an infinitely short time is to produce a change in velocity (or momentum) but no displacement of the particle. is just as if the force had no time to produce a displacement in the body, even though it produces a change in the momentum. shall call such forces momentary or impulsive forces. ure is the quantity J or the amount of change in momentum they Though such forces are strictly ideal, they are approxiproduce. mated in nature by such impulses as the blow of a hammer and the explosion of the powder back of a projectile. The reader should be able to show very simply that when computing the effect of an impulsive force on a particle all finite forces acting during the same interval may be neglected.

Now let us ask ourselves about the cumulative effect of a force over the space traveled by the particle acted on by the force. In analogy with the measure of the time cumulative effect just discussed above, it is natural to take as the space cumulative effect the space integral of the force. But since this involves the product of two vectors it is necessary first to define what we shall mean by this.

Suppose we have two vectors A and B, written in terms of their rectangular components as follows [cf. (1·3-6)]:

$$A = iA_x + jA_y + kA_z,$$

$$B = iB_x + jB_y + kB_z.$$
(1·10-4)

We shall define the so-called dot or scalar product of A and B as

$$\mathbf{A} \cdot \mathbf{B} = A_x B_x + A_y B_y + A_z B_z. \tag{1.10-5}$$

It is clear that the commutative law holds. Thus

IG CHO

$$\mathbf{A} \cdot \mathbf{B} = \mathbf{B} \cdot \mathbf{A}. \tag{1.10-6}$$

ment can ever be devised to test them fully and unambiguously. They may be taken to represent our idea as to the most general and simple basis, compatible with experiment, for the development of mechanics. Beginning with them as assumptions we shall make various deductions which may then be checked by the results of the most careful experiments which can be carried out. We shall see in the course of our work that beside the three laws of Newton, other more compact formulations of the principles of mechanics are possible and very valuable for certain purposes.

1.10. Impulse and Kinetic Energy. Work and Power. In addition to mass and force there are certain other mechanical concepts so important that it seems desirable to introduce them here, even though we shall postpone their detailed application to Chapter IV.

What is the cumulative effect of a force acting on a particle as time passes? The natural measure of this is the time integral of the force over the interval in question. This is known as the *impulse* of the force in this interval. A simple but significant theorem can be proved about this quantity. Thus if we employ the equation of motion (1.7–5) we can write

$$\int_{t_0}^{t_1} \mathbf{F} dt = m \int_{t_0}^{t_1} \frac{d\mathbf{v}}{dt} dt = m \int_{\mathbf{v}_0}^{\mathbf{v}_1} d\mathbf{v}, \qquad (1.10-1)$$

where v_0 and v_1 are the velocities of the particle at the instants t_0 and t_1 respectively. Integration yields

$$\int_{t_0}^{t_1} \mathbf{F} \, dt = m \mathbf{v}_1 - m \mathbf{v}_0. \tag{1.10-2}$$

It thus appears that the impulse of the resultant force is the *change* in *momentum* of the particle during the interval in which the force acts. This is known as the impulse-momentum theorem. It is of particular interest for forces which are of short duration.

For example, though the force may be very great (i.e., mathematically approach infinity) the time integral (1·10–2) may nevertheless be finite if the time interval is small enough. Thus we may have for $\mathbf{F} \doteq \infty$ and $\tau \doteq 0$ for a particle of mass m

$$m(\mathbf{v}_2 - \mathbf{v}_1) = \int_0^{\tau} \mathbf{F} dt = \mathbf{J},$$
 (1.10-3)

where v₁ and v₂ are the velocities of the particle at the beginning and end of the time interval τ and J is a finite vector quantity. A very interesting though curious feature of this problem is seen in the following. Suppose the velocity of the particle remains always finite and let v_{max} be its largest value during the interval. Then the actual distance traveled during the duration of application of the force is certainly less than τv_{max} . But as $\tau \doteq 0$ this approaches zero, and hence we conclude that the effect of this infinite force for an infinitely short time is to produce a change in velocity (or momentum) but no displacement of the particle. is just as if the force had no time to produce a displacement in the body, even though it produces a change in the momentum. shall call such forces momentary or impulsive forces. ure is the quantity J or the amount of change in momentum they Though such forces are strictly ideal, they are approxiproduce. mated in nature by such impulses as the blow of a hammer and the explosion of the powder back of a projectile. The reader should be able to show very simply that when computing the effect of an impulsive force on a particle all finite forces acting during the same interval may be neglected.

Now let us ask ourselves about the cumulative effect of a force over the space traveled by the particle acted on by the force. In analogy with the measure of the time cumulative effect just discussed above, it is natural to take as the space cumulative effect the space integral of the force. But since this involves the product of two vectors it is necessary first to define what we shall mean by this.

Suppose we have two vectors A and B, written in terms of their rectangular components as follows [cf. (1·3-6)]:

$$\mathbf{A} = \mathbf{i}A_x + \mathbf{j}A_y + \mathbf{k}A_z,$$

$$\mathbf{B} = \mathbf{i}B_x + \mathbf{j}B_y + \mathbf{k}B_z.$$
(1·10-4)

We shall define the so-called dot or scalar product of A and B as

$$\mathbf{A} \cdot \mathbf{B} = A_x B_x + A_y B_y + A_z B_z. \tag{1.10-5}$$

It is clear that the commutative law holds. Thus

IG CHO

$$\mathbf{A} \cdot \mathbf{B} = \mathbf{B} \cdot \mathbf{A}. \tag{1.10-6}$$

24 THE ELEMENTAL CONCEPTS OF MECHANICS

ment can ever be devised to test them fully and unambiguously. They may be taken to represent our idea as to the most general and simple basis, compatible with experiment, for the development of mechanics. Beginning with them as assumptions we shall make various deductions which may then be checked by the results of the most careful experiments which can be carried out. We shall see in the course of our work that beside the three laws of Newton, other more compact formulations of the principles of mechanics are possible and very valuable for certain purposes.

1.10. Impulse and Kinetic Energy. Work and Power. In addition to mass and force there are certain other mechanical concepts so important that it seems desirable to introduce them here, even though we shall postpone their detailed application to Chapter IV.

What is the cumulative effect of a force acting on a particle as time passes? The natural measure of this is the time integral of the force over the interval in question. This is known as the *impulse* of the force in this interval. A simple but significant theorem can be proved about this quantity. Thus if we employ the equation of motion (1.7–5) we can write

$$\int_{t_0}^{t_1} \mathbf{F} dt = m \int_{t_0}^{t_1} \frac{d\mathbf{v}}{dt} dt = m \int_{\mathbf{v}_0}^{\mathbf{v}_1} d\mathbf{v}, \qquad (1.10-1)$$

where \mathbf{v}_0 and \mathbf{v}_1 are the velocities of the particle at the instants t_0 and t_1 respectively. Integration yields

$$\int_{t_0}^{t_1} \mathbf{F} dt = m\mathbf{v}_1 - m\mathbf{v}_0. \tag{1.10-2}$$

It thus appears that the impulse of the resultant force is the *change* in *momentum* of the particle during the interval in which the force acts. This is known as the impulse-momentum theorem. It is of particular interest for forces which are of short duration.

For example, though the force may be very great (i.e., mathematically approach infinity) the time integral (1·10–2) may nevertheless be finite if the time interval is small enough. Thus we may have for $\mathbf{F} \doteq \infty$ and $\tau \doteq 0$ for a particle of mass m

$$m(\mathbf{v}_2 - \mathbf{v}_1) = \int_0^{\tau} \mathbf{F} dt = \mathbf{J},$$
 (1.10-3)

where \mathbf{v}_1 and \mathbf{v}_2 are the velocities of the particle at the beginning and end of the time interval τ and J is a finite vector quantity. A very interesting though curious feature of this problem is seen in the following. Suppose the velocity of the particle remains always finite and let v_{max} be its largest value during the interval. Then the actual distance traveled during the duration of application of the force is certainly less than τv_{max} . But as $\tau \doteq 0$ this approaches zero, and hence we conclude that the effect of this infinite force for an infinitely short time is to produce a change in velocity (or momentum) but no displacement of the particle. is just as if the force had no time to produce a displacement in the body, even though it produces a change in the momentum. shall call such forces momentary or impulsive forces. ure is the quantity I or the amount of change in momentum they Though such forces are strictly ideal, they are approxiproduce. mated in nature by such impulses as the blow of a hammer and the explosion of the powder back of a projectile. The reader should be able to show very simply that when computing the effect of an impulsive force on a particle all finite forces acting during the same interval may be neglected.

Now let us ask ourselves about the cumulative effect of a force over the space traveled by the particle acted on by the force. In analogy with the measure of the time cumulative effect just discussed above, it is natural to take as the space cumulative effect the space integral of the force. But since this involves the product of two vectors it is necessary first to define what we shall mean by this.

Suppose we have two vectors A and B, written in terms of their rectangular components as follows [cf. (1.3-6)]:

$$\mathbf{A} = \mathbf{i}A_x + \mathbf{j}A_y + \mathbf{k}A_z,$$

$$\mathbf{B} = \mathbf{i}B_x + \mathbf{j}B_y + \mathbf{k}B_z.$$
(1·10-4)

We shall define the so-called dot or scalar product of A and B as

$$\mathbf{A} \cdot \mathbf{B} = A_x B_x + A_y B_y + A_z B_z. \tag{1.10-5}$$

It is clear that the commutative law holds. Thus

SIG CHA

$$\mathbf{A} \cdot \mathbf{B} = \mathbf{B} \cdot \mathbf{A}. \tag{1.10-6}$$

Moreover if A, B, C are any three vectors, the distributive law holds, viz.:

$$(A + B) \cdot C = A \cdot C + B \cdot C. \qquad (1.10-7)$$

The proof is left to the reader. Since the magnitudes of A and B are given by (cf. Sec. 1.3 again)

$$A = |\mathbf{A}| = \sqrt{A_x^2 + A_y^2 + A_z^2},$$

$$B = |\mathbf{B}| = \sqrt{B_x^2 + B_y^2 + B_z^2},$$
(1·10-8)

we may write (1.10-5) in the form

$$\mathbf{A} \cdot \mathbf{B} = AB \left(\frac{A_z B_x}{AB} + \frac{A_y B_y}{AB} + \frac{A_z B_z}{AB} \right) \cdot \tag{1.10-9}$$

But

$$A_x = A \cos \alpha_1, \quad A_y = A \cos \alpha_2, \quad A_z = A \cos \alpha_3, B_z = B \cos \beta_1, \quad B_y = B \cos \beta_2, \quad B_z = B \cos \beta_3,$$
 (1·10-10)

where $\cos \alpha_1$, $\cos \alpha_2$, $\cos \alpha_3$ are the direction cosines of A (cf. Sec. 1·2), and $\cos \beta_1$, $\cos \beta_2$, $\cos \beta_3$ are the direction cosines of B. A fundamental theorem of analytic geometry gives

$$\cos \theta = \cos \alpha_1 \cos \beta_1 + \cos \alpha_2 \cos \beta_2 + \cos \alpha_3 \cos \beta_3 \quad (1.10-11)$$

as the cosine of the angle between the vectors $\bf A$ and $\bf B$. Hence $(1\cdot 10-9)$ becomes

$$\mathbf{A} \cdot \mathbf{B} = AB \cos \theta. \tag{1.10-12}$$

This result is indeed sometimes used as the definition of the scalar product of the vectors A and B.

By the use of $(1\cdot10-12)$ we readily show the following properties of the unit vectors \mathbf{i} , \mathbf{j} , \mathbf{k} .

$$\mathbf{i} \cdot \mathbf{i} = \mathbf{j} \cdot \mathbf{j} = \mathbf{k} \cdot \mathbf{k} = 1,$$

 $\mathbf{i} \cdot \mathbf{j} = \mathbf{j} \cdot \mathbf{k} = \mathbf{k} \cdot \mathbf{i} = 0.$ (1·10-13)

Reverting to the definition (1·10-5) we can assume that **A** and **B** are functions of the time and differentiate, obtaining

$$\begin{split} \frac{d}{dt} \left(\mathbf{A} \cdot \mathbf{B} \right) &= \dot{A}_x B_x + A_x \dot{B}_x + \dot{A}_y B_y + A_y \dot{B}_y \\ &+ \dot{A}_z B_z + A_z \dot{B}_z \\ &= \dot{\mathbf{A}} \cdot \mathbf{B} + \mathbf{A} \cdot \dot{\mathbf{B}}, \end{split} \tag{1.10-14}$$

where

$$\dot{\mathbf{A}} = \mathbf{i}\dot{A}_x + \mathbf{j}\dot{A}_y + \mathbf{k}\dot{A}_z, \qquad (1.10-15)$$

it being assumed that i, j, k remain fixed as time passes.

We are now ready to consider the space integral of the force. We shall define it as follows:

$$\int_{\mathbf{r}_0}^{\mathbf{r}_1} \mathbf{F} \cdot d\mathbf{r}, \qquad (1.10-16)$$

where \mathbf{r}_0 is the position vector of the particle at the beginning of its motion, and \mathbf{r}_1 is the position vector at the end of the motion. The infinitesimal displacement of the particle in the neighborhood of position vector \mathbf{r} is denoted by $d\mathbf{r}$. The integrand is the space effect of the force during the displacement $d\mathbf{r}$, when the force has the (nearly) instantaneous value \mathbf{F} . The whole integral is then taken as the space cumulative effect of the force during the motion from \mathbf{r}_0 to \mathbf{r}_1 . We shall call the integral in (1·10–16) the work done by the force on the particle during its displacement and denote it by W. Let us evaluate W for the important case in which the force is the resultant force acting on the particle. Then we may replace \mathbf{F} by $m\ddot{\mathbf{r}}$ from the equation of motion (1·7–4) and get

$$W = \int_{r_0}^{r_1} m\ddot{\mathbf{r}} \cdot d\mathbf{r} = \int_{r_0}^{r_1} m\ddot{\mathbf{r}} \cdot \dot{\mathbf{r}} \, dt, \qquad (1.10-17)$$

where we have replaced $d\mathbf{r}$ by $\mathbf{\dot{r}}$ dt. Let us examine $\mathbf{\ddot{r}} \cdot \mathbf{\dot{r}}$. If we differentiate $\mathbf{\dot{r}} \cdot \mathbf{\dot{r}}$ with respect to time, we get from $(1\cdot10-14)$

$$\frac{d}{dt}(\dot{\mathbf{r}}\cdot\dot{\mathbf{r}}) = 2\ddot{\mathbf{r}}\cdot\dot{\mathbf{r}},\tag{1.10-18}$$

recalling that the scalar product is commutative. Hence (1·10-17) becomes

$$W = \frac{m}{2} \int d(\dot{\mathbf{r}} \cdot \dot{\mathbf{r}}) = \frac{m}{2} \int_{v_0}^{v_1} d(v^2), \qquad (1.10-19)$$

where we have replaced $\dot{\mathbf{r}} \cdot \dot{\mathbf{r}}$ by v^2 , since

$$v = \dot{r}$$

and have changed our limits to the initial and final speed values

respectively. Now the integration in (1·10-19) can be readily carried out and yields

$$W = \frac{1}{2}mv_1^2 - \frac{1}{2}mv_0^2. \tag{1.10-20}$$

This reminds us of the impulse-momentum theorem $(1\cdot10-2)$ in that the work done by the resultant force appears as the difference between the values of a certain function of mass and velocity at the beginning and the end of the motion. The function here is $\frac{1}{2}mv^2$, a scalar quantity which was known in the 18th century as the vis viva of the particle but which during the 19th century came to be called the kinetic energy. We shall, in general, denote it by the symbol K. The theorem we have just proved then says that the work done by the resultant force acting on a particle is equal to the change in kinetic energy experienced. We call this result the work-kinetic energy theorem. We must be careful to note that it applies only to the work done by the resultant force. Actually a force can act on a particle and do no work on it at all. Thus to revert to $(1\cdot10-16)$ we see that if F is at right angles to the displacement $d\mathbf{r}$, we have

$$dW = \mathbf{F} \cdot d\mathbf{r} = 0$$
.

In other words, whenever a force is perpendicular to the displacement of the particle its space cumulative effect is zero or it does no work. This, for example, is the case for the centripetal force discussed in Sec. 1.5. In general

$$dW = \mathbf{F} \, d\mathbf{r} \cos \theta, \qquad (1.10-21)$$

where θ is the angle between **F** and $d\mathbf{r}$.

We must now consider the matter of units. The work done when a particle is moved one centimeter while acted upon by a force of one dyne in the direction of the motion is called an erg. Because of the smallness of this quantity, a larger unit is used. This is equal to 10^7 ergs and is called the joule. The student should verify that $\frac{1}{2}mv^2$ has the same dimensions and therefore the same units as work; i.e., dimensionally, kinetic energy and work are the same, as naturally they should be by definition. The absolute English unit of work and energy is the foot poundal; this is the work done when a particle is moved through a distance of one foot by a force of one poundal in the direction of the motion. Engineers never use this unit; they prefer the foot pound, which is the work

done in the motion of a particle through one foot under the action of a pound of force (lb. wt.). From the relation between the two units of force we note that 32 foot poundals are equal to one foot pound. In the m.k.s. system, mentioned in Sec. 1.8, the natural unit of work is the Newton meter. But a moment's reflection shows us that the Newton meter is exactly equal to the joule. This indicates one advantage of the m.k.s. system of units.

The rate at which work is done is called *power*. Denoting the latter by P we have the general relation

$$P = \dot{W}. \tag{1.10-22}$$

The metric unit (which is the same here as the m.k.s. unit) is the watt or joule/sec, the kilowatt being 1000 watts. In the English system the corresponding unit foot pound/sec has no name, but 550 foot pounds/sec is called one horsepower (hp.). The student is advised to note that the kilowatt hour and the horsepower hour are not units of power but units of work, equivalent respectively to 3.60×10^6 joules and 1.98×10^6 foot pounds.

By dividing by dt both sides of the relation

$$dW = \mathbf{F} \cdot d\mathbf{r}$$

we obtain

$$P = \mathbf{F} \cdot \dot{\mathbf{r}} = \mathbf{F} \cdot \mathbf{v}, \qquad (1.10-23)$$

which is often a useful relation.

1.11. Résumé. With the ideas of the preceding sections in mind we are now prepared to take up some concrete illustrations which will develop our understanding of the laws of mechanics. In the next two chapters we shall discuss the motion of a particle The main problem we shall meet there under various conditions. will be the setting up and solution of the equation (or equations) of motion, which is only the mathematical expression of the second law, namely (1.7-4). The physics of the problem enters with the significant choice of the force F and the appropriate initial conditions of the motion, i.e., the so-called "boundary conditions," the initial or final position and velocity of the particle. The rest of the problem will be mathematical. But each illustration will be so arranged that the results will tie up intimately with the experience of our actual world. This will strengthen our grasp of the fundation mental concepts already introduced in the present chapter and

afford a preparation for the more elaborate treatment to follow. In Chapter IV we shall show how the concept of energy can be used to particular advantage in the solution of mechanical problems.

There is one point about the use of the equation of motion which merits a general remark before the solution of specific problems is undertaken. We have already noted in Sec. 1.2 that position is relative and the same must be true of displacement, velocity and acceleration. It will therefore naturally occur to the thoughtful student to inquire whether the equation of motion (1.7-4) will retain its present simple form in all systems of reference. answer to this is in the negative. The most general system in which the equation applies is that fixed relatively to the average position of the fixed stars. Nevertheless, it is easy to show that the equation retains the same form in any system which moves with constant velocity (without rotation) with respect to the above mentioned system. For suppose x, y, z are the coördinates of a particle in the first reference system, while x', y', z' are the coordinates in a system moving in the x direction with velocity v relative to the first. We then have the following relations connecting the coordinates in the two systems

$$x' = x - vt - a,$$

 $y' = y - b,$
 $z' = z - c,$ (1.11-1)

where a, b, c are constants giving the position of the origin of the second system relative to the first when t = 0. But it then follows at once that

$$\ddot{x}' = \ddot{x},$$

 $\ddot{y}' = \ddot{y},$
 $\ddot{z}' = \ddot{z},$
(1·11-2)

whence the accelerations are the same in the two systems. It must, of course, be noted that this result would *not* be true were the two systems *accelerated* relatively to each other.

Our discussion has, to be sure, neglected the effect of the transformation on the force side of the equation. But we recall from Sec. 1.7 that **F** is a function of the relative position coördinates of the interacting particles. Relative position coördinates are not altered by the transformation (1.11–1).

PROBLEMS

- 1. Determine the expression for the distance between two points in a plane (a) in terms of rectangular coördinates, (b) in terms of polar coördinates, (c) in terms of oblique coördinates with the axes making an angle of 60° with each other.
- 2. The position of a point in a plane with respect to rectangular coördinate axes is given by (x, y). If the axes are rotated through angle ϕ , what are the new position coördinates?
- 3. The position of a point P in space with respect to rectangular coördinate axes is given by (x, y, z). If the axes are rotated with the origin held fixed, the position of the original point P is now given by (x', y', z'). Find the expressions for x', y', z' in terms of x, y, z and the direction cosines of the new axes with respect to the old axes.
- 4. Prove that the distance between any two points in three-dimensional space remains invariant under a rotation of axes as that discussed in Problem 3.
- 5. A flywheel of radius 1 meter starts from rest with a constant angular acceleration of 1 rev/sec². Find the resultant acceleration of a particle on the periphery at the end of 1 second.
- 6. Compare the instantaneous linear velocities relative to the ground of the hub of a wheel and the highest point of the rim.
- 7. A ship sails southeast at 15 mi/hr and the wind blows from the south at 10 mi/hr. In what position will a masthead pennant come to rest?
- 8. A boat is headed southwest with a speed of 10 knots, while another boat is headed due east with a speed of 15 knots. Find the relative velocity of the boats.
- 9. Two particles move with equal speeds v in opposite directions in the same circle. In what position will their relative velocity have a maximum magnitude and what will it be? What is the magnitude of the relative velocity when the position vectors of the two particles make the angle 45° with each other? Draw a diagram showing the actual relative velocities.
- 10. Particle A moves with a uniform velocity of 100 cm/sec in a direction making angles of 60° and 45° with the x and y axes respectively. Particle B has a constant velocity of 50 cm/sec along the z axis. Find the components along the x, y, and z axes respectively of the difference between the two velocities.
- 11. Show by a simple example that the resultant of two *finite* angular displacements depends in general on the order in which the displacements are performed. Are there exceptions? Show by a construction why it is plausible that the resultant of two *infinitesimal* angular displacements is independent of the order in which such displacements are performed.
- 12. On the simple Bohr theory of the structure of the hydrogen atom a negatively charged electron revolves in a circular orbit about a positively charged nucleus. Denoting the charge on the electron by -e, that on the

nucleus by +e, and the mass of the electron by m, find the expression for the angular velocity ω of the electron corresponding to orbital radius a. Calculate ω for $a=0.53\times 10^{-8}$ cm, $m=9\times 10^{-28}$ grams, and $e=4.80\times 10^{-10}$ electrostatic units. (Hint: Use the electrostatic Coulomb law of force between point charges.)

- 13. A particle moves with uniform angular velocity in a circle of radius a. Derive the expressions for the rectangular components $(\dot{x}$ and $\dot{y})$ of the peripheral velocity in terms of the angular displacement θ from the x axis as polar axis. Find the similar expressions for \ddot{x} and \ddot{y} .
- 14. A particle moves in a plane with respect to fixed rectangular axes. Express its velocity and acceleration components relative to a set of rectangular axes rotating about the same origin with constant angular velocity ω in terms of the corresponding components relative to the fixed axes. Specialize to the case where the particle is at rest in the fixed system.
- 15. A particle of mass 100 grams and traveling along the x axis in the direction of increasing x with the constant speed 100 cm/sec collides at the origin with a particle of mass 200 grams traveling along the y axis in the direction of increasing y with the constant speed 150 cm/sec. If the two particles lock together after the collision, find the velocity in direction and magnitude of the combined particles. Compare the kinetic energy after collision with that of the individual particles before collision.
- 16. Two particles of masses m_1 and m_2 respectively suffer acceleration due to their mutual interaction. Find the expression for their relative acceleration in terms of the actual acceleration of either particle in a fixed reference system. What does it become when either mass is very large compared with the other?
- 17. In Problem 12, derive the expression for the kinetic energy of the electron moving in the circular orbit of radius a.
- 18. A force $F = kte^{-at}$, where k and a are positive constants, acts on a particle of mass m constrained to move in the x direction. Find the change in momentum produced by the force in time t. If the particle starts from rest, how far will it go in a time t_0 equal to that at which the force attains its maximum value?
- 19. In Problem 18, if the particle starts from rest, find the kinetic energy attained in time t_0 . Find the maximum kinetic energy attained and the time at which it is attained.
- 20. In Problem 18, if the particle starts from rest, find the instantaneous power at time t. Also find the average power over the time interval τ .

CHAPTER II

RECTILINEAR MOTION OF A PARTICLE

2.1. Free Motion and the Uniform Field. As an introduction to the study of the motion of a particle, let us consider motion along a straight line. The simplest type of this is, of course, "free" motion, that is, motion without acceleration, i.e., under the influence of no force. The velocity is constant and the distance traveled in time t, assuming that the motion takes place along the x axis, is $x - x_0$, where

$$x = vt + x_0. (2.1-1)$$

The velocity is v in magnitude and the original displacement of the particle at t = 0 is x_0 .

From this trivial case we pass at once to the more important one in which the particle moves with a constant acceleration. We shall call this motion in a uniform (i.e., constant) field of force. If the mass of the particle is m, the equation of motion (1.7-4) when reduced to scalar form becomes

$$m\ddot{x} = F_0, \qquad (2\cdot 1-2)$$

where F_0 represents the constant force. The equation (2·1-2) may be written in the form $\dot{v} = F_0/m$, whence $dv = F_0/m \cdot dt$. Integrating both sides we get

$$v = \dot{x} = \int \frac{F_0}{m} dt = \frac{F_0}{m} t + c,$$

where c is a constant of integration. In value it is equal to v when t = 0, i.e., the initial velocity, which we may call v_0 . So we have

$$v = \frac{F_0}{m}t + v_0. {(2.1-3)}$$

A second integration in similar fashion gives the distance from the

¹ The motion being rectilinear, we are interested in its scalar magnitude only.

origin at time t, viz.,

$$x = \frac{1}{2} \left(\frac{F_0}{m} \right) t^2 + v_0 t + x_0, \tag{2.1-4}$$

where we have already replaced the constant of integration by x_0 , the initial displacement of the particle. This simple illustration shows clearly that in order to use for purposes of computation the integrated equation of motion, we must know the particle's displacement and velocity at some chosen instant (here t=0). These are, so to speak, the boundary conditions of the motion, though they are usually referred to as the initial conditions.

By elimination of the time t between eqs. $(2\cdot 1-3)$ and $(2\cdot 1-4)$ the reader may show the existence of the following important relation between velocity and distance traveled, viz.,

$$v^{2} = \frac{2F_{0}}{m} (x - x_{0}) + v_{0}^{2}.$$
 (2-1-5)

The most important illustration of this type of motion is furnished by a particle falling freely under the influence of gravity near the earth's surface, neglecting the air resistance. Then F_0 becomes the weight of the particle, which is numerically equal to mg, where g is the constant acceleration of gravity. We are here considering the *upward* direction as positive, while the weight is a downward force. Hence $F_0/m = -g$ and $(2\cdot 1-4)$ becomes

$$x = -\frac{1}{2}gt^2 + v_0t + x_0, \qquad (2.1-6)$$

and $(2\cdot1-5)$ is now

$$v^2 = -2g(x - x_0) + v_0^2. (2.1-7)$$

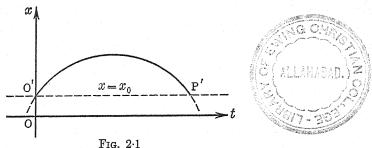
If we wish to discuss the height which a particle thrown into the air with initial upward velocity v_0 will attain, we merely substitute v = 0 in eq. (2·1-7), whence the height is given at once by

$$h = \frac{v_0^2}{2g} \cdot \tag{2.1-8}$$

The same result follows from the application of eqs. $(2\cdot1-3)$ and $(2\cdot1-6)$. Incidentally the reader may easily show that the particle spends as long a time going up as it does coming down and

¹ We also neglect the small effect due to the rotation of the earth. See end of Sec. 3-3. Also Sec. 7-12.

that it reaches the ground with a downward velocity equal in magnitude to its original upward velocity. These facts are very neatly brought out by the construction of a space-time graph, plotting the coördinate x against the time t, by means of the equation $(2\cdot 1-6)$. The result, of course, is a parabola as shown in Fig. 2·1, where the axis of abscissas is the time axis and the ordinates are values of the upward displacement. What may be called the space-time history of the particle from the instant it is projected upward to the time when it strikes the ground is completely given by the portion of the curve between O' and P'.



Moreover, the *slope* of the curve at each point yields the *instantaneous velocity* of the particle at the corresponding time and place. We shall find this type of graph of value in the further study of problems in motion.

It has already been pointed out that in the above discussion we are neglecting the effect of air resistance, in order to simplify the problem. (See Sec. 2.6 for the more complicated case.)

There are some important examples of rectilinear motion under the influence of gravity in which the particle is "constrained" to move in a certain direction or with an acceleration which while constant is different from "g." Such cases logically belong to the discussion in Chapter VIII. But the physical ideas involved are so common and so important that we shall discuss here the problem of motion along an inclined plane and the Atwood machine.

Consulting Fig. 2-2, which represents a plane inclined at an angle θ to the horizontal, let a particle of mass m move without friction on this plane. The force of gravity is mg, but this acts vertically downward while the particle is constrained to move along the plane. The net, unbalanced force F which acts in the

direction of the motion has therefore the magnitude $mg \sin \theta$, and the equation of motion is

 $m\ddot{x} = mg\sin\theta, \qquad (2\cdot1-9)$

where x is measured positively down the plane, for convenience. The results of the first part of this section will then follow in this

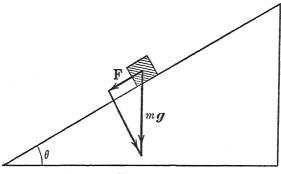


Fig. 2.2

case with the substitution of $g \sin \theta$ for g. For example, the application of eq. $(2\cdot 1-7)$ reveals the interesting fact that since

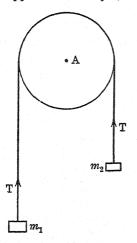


Fig. 2.3

 $l\sin\theta=h$, where l is the length and h the height of the plane, the velocity along the plane acquired at the foot of the plane is equal to the velocity which would be acquired were the particle to fall freely through a distance equal to the height of the plane. The times of fall for the same initial velocities are, of course, different. In fact, the reader should show that the time for descent along the plane is to that down the height of the plane in the ratio $1:\sin\theta$.

In the Atwood machine (shown diagrammatically in Fig. 2.3) there are two particles of mass m_1 and m_2 respectively at the ends of a flexible inextensible string wrapped around a frictionless peg. A.

Suppose $m_1 > m_2$. We proceed to set up an equation of motion for each particle by determining the unbalanced force on each. The string pulls up on each mass particle with a force T, the so-called "tension" in the string, which we shall assume is not

altered when the latter passes over a perfectly smooth peg. The equations of motion therefore are

$$m_1g - T = m_1\ddot{x}_1$$
 $m_2g - T = m_2\ddot{x}_2$
 $(2\cdot 1-10)$

where x_1 and x_2 are the corresponding displacements of m_1 and m_2 measured downward. Now since the one mass rises with the same acceleration with which the other falls, we have

$$\ddot{x}_1 = -\ddot{x}_2 = \ddot{x},$$
 (2·1-11)

and hence by subtraction of the two equations in (2·1–10) there results

 $\ddot{x} = \frac{m_1 - m_2}{m_1 + m_2} \cdot g, \tag{2.1-12}$

and the solution follows as in the previous cases. The tension in the string is evaluated by substitution from $(2\cdot 1-12)$ into either of the equations $(2\cdot 1-10)$. The utility of this device in studying the acceleration of gravity is evident from the fact that the actual acceleration of either particle may be made as small as desirable by choosing m_1 and m_2 sufficiently close together in magnitude.

2.2. Motion in a Field Proportional to the First Power of the Distance. The next simplest type of rectilinear motion is perhaps that in which the force is directly proportional to the distance from some fixed point in the straight line in which the motion takes place, a state of affairs not uncommon in physical phenomena. In such a case the equation of motion has the general form

$$m\ddot{x} = cx, \qquad (2\cdot 2-1)$$

where c is a constant. Two cases will arise, according as c is negative or positive. The former is by far the more important for our purpose and will be considered first. Let c=-k, where k is positive. Probably the simplest way of solving the different

¹ This book is not primarily a mathematical treatise. We shall therefore attempt to develop and solve each problem by the simplest mathematical methods available. This explains why, in the present case, we do not enter into the more elegant mathematical methods for the solution of differential equations of the type of (2·2–1). These will be found in any fairly comprehensive text on differential equations (e.g., A. Cohen, "Elementary Treatise on Differential Equations," Heath, N. Y., 1906, Chap. VII), and will of course be of great interest and ultimate use to the thoughtful student of theoretical physics.

tial equation (2·2–1) is to multiply both sides by \dot{x} dt. Thus we have

$$m\dot{x}\ddot{x} dt = -kx dx.$$

But this is simply

$$\frac{1}{2}md(\dot{x})^2 = -\frac{1}{2}kd(x)^2. \tag{2.2-2}$$

Integrating once then yields

$$\frac{1}{2}m\dot{x}^2 = -\frac{1}{2}kx^2 + C_1, \qquad (2\cdot 2-3)$$

where C_1 is a constant of integration. This so-called first integral of the equation of motion has a particularly significant form, which will be emphasized in Chapter IV. The value of C_1 will depend on the initial conditions imposed on the motion. It must be emphasized that without these conditions we could not obtain specific physical results from our integration. The importance of such conditions will be continually evident throughout the text. Let us assume that when x=R, $v=\dot{x}=0$. Then $C_1=\frac{1}{2}kR^2$, and eq. (2·2-3) can be made to take the form

$$\dot{x} = \pm \sqrt{\frac{k}{m}} \cdot \sqrt{R^2 - x^2}, \qquad (2.2-4)$$

which yields the velocity of the particle for any distance x from the chosen origin, subject, of course, to the above condition, which is arbitrary in nature. However, it should be noted from eq. $(2\cdot2-3)$ that the velocity must be zero for *some* value of x. We have merely specified this value as R in our above assumption. Separating variables in eq. $(2\cdot2-4)$ gives us

$$\frac{dx}{\pm\sqrt{R^2-x^2}}=\sqrt{\frac{k}{m}}\,dt,$$
 (2.2-5)

which on integration becomes

$$\operatorname{arc} \frac{\sin}{\cos} \left\{ \left(\frac{x}{R} \right) = \sqrt{\frac{k}{m}} t + C_2, \qquad (2.2-6) \right\}$$

where the notation of the left-hand side indicates that arc sin (x/R) or arc cos (x/R) results according as we choose the plus or minus sign in $(2\cdot2-5)$. C_2 is a second constant of integration, the value of which must be determined from a second initial condition giving the initial value of the displacement x. Thus if

x=R when t=0, we have $C_2=\arctan\left\{\begin{array}{l} \sin \\ \cos \end{array}\right\}$ (1), which equals either $\pi/2$ or zero, according as we use the sine or cosine function. On the other hand, if x=0 for t=0, $C_2=0$ or $\pi/2$, according as the sine or cosine is used.

The explicit expression for the displacement x in terms of the time is then

$$x = R \frac{\sin}{\cos} \left(\sqrt{\frac{k}{m}} \cdot t + C_2 \right) \cdot \tag{2-2-7}$$

The space-time graph of this case is the simple sine or cosine curve familiar to the student from elementary mathematics, and

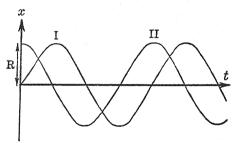


Fig. 2.4

the motion is known by the term simple harmonic. If we choose $C_2 = 0$, the result is shown in Fig. 2.4, where curve I corresponds to the sine and II to the cosine. The two can, of course, be made to coincide by a relative displacement along the t axis. In simple harmonic motion there is thus a maximum displacement of the particle from the chosen origin and its value is R, defined as the amplitude of the motion. The motion repeats itself, i.e., is oscillatory or periodic, and the time for a complete oscillation, denoted by P and called the period, is given by

$$P = 2\pi \sqrt{\frac{m}{k}}.$$
 (2.2–8)

This follows from an inspection of eq. $(2\cdot2-7)$. Thus to consider the sine term only, $x=R\sin C_2$ for t=0. The same value of x will next recur at a time P for which $\sin (\sqrt{k/m} P + C_2) = \sin C_2$. But this will happen for $\sqrt{k/m} P = 2\pi$, whence $(2\cdot2-8)$ follows. The reciprocal of P is called the *frequency* and is represented by ν .

It is the number of complete oscillations per second. The displacement may then be written in the simplified form (using here the cosine only)

 $x = R\cos(2\pi\nu t + \epsilon), \qquad (2\cdot 2-9)$

where C_2 has been replaced by ϵ . The argument of the cosine, viz., $2\pi\nu t + \epsilon$, is called the *phase* of the simple harmonic motion and ϵ , its value for t = 0, is known as the *epoch*.

From the above discussion the reader will see at once that in simple harmonic motion the particle has its maximum velocity when x=0 (i.e., while passing through the chosen origin) while its minimum velocity (viz. zero) occurs at $x=\pm R$, i.e., at the two ends of its path. On the other hand, the maximum acceleration or force corresponds to x=R, while the minimum (zero in value) occurs at x=0.

The reader will have noted that the treatment of simple harmonic motion given immediately above is purely analytic in nature. That is, we began with a differential equation, i.e., the equation of motion, and by two successive integrations attained the expressions for the velocity and displacement which were rendered specific by the introduction of the initial or boundary conditions. This is the straightforward method of discussing the problem. Yet in many cases it is illuminating to start at the other end, i.e., to commence with a certain well-defined type of

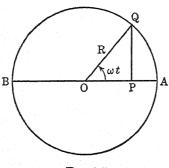


Fig. 2.5

motion which appears to possess features of interest and then deduce its properties. This method of approach to a dynamical problem is very often the simplest. Let us now apply it to the case of simple harmonic motion. We shall suppose that a particle Q moves with uniform angular velocity ω rad/sec about a circle with radius R and center at O (see Fig. 2-5). Let us consider the motion of the projection of Q

on any straight line in the plane of the circle. For simplicity choose this straight line as the diameter AB and let P be the projection in question. As the particle Q moves about the circle, P

moves back and forth on the line AB with a maximum excursion from O of magnitude R. If the motion is assumed to begin at A when t=0, the displacement of P from O at any subsequent instant t is, of course,

$$x = R \cos \omega t. \tag{2.2-10}$$

Successive differentiations with respect to the time yield the instantaneous velocity and acceleration respectively, namely,

$$v = \dot{x} = -R\omega \sin \omega t$$

$$= \pm R\omega \sqrt{1 - \frac{x^2}{R^2}}$$

$$= \pm \omega \sqrt{R^2 - x^2}, \qquad (2.2-11)$$

and

$$a = \ddot{x} = -R\omega^2 \cos \omega t$$
$$= -\omega^2 x. \tag{2.2-12}$$

If the number of revolutions per second of Q is $\nu = \omega/2\pi$, the frequency of the motion of P will likewise be ν . Comparison then shows that eq. $(2\cdot 2-12)$ is equivalent to $(2\cdot 2-1)$ with $k/m = \omega^2 =$

 $4\pi^2v^2$. Likewise eq. (2·2–11) is identical with (2·2–4). In other words, the motion as defined above kinematically is precisely that which results when a particle moves in a straight line in a field of force varying directly as the distance from a fixed point on the line and everywhere opposite in direction to the displacement, and hence it is by definition simple harmonic motion. In the two treatments we have thus secured both a dynamical and a kinematic view of this important type of motion.

It is now in order to discuss an illustration. We shall consider the case of a spring (assumed here for simplicity as

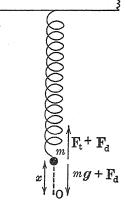


Fig. 2.6

massless) one end of which is fastened to a rigid ceiling and to the other end of which there is attached a heavy particle of mass m (Fig. 2-6). At the moment when we first consider the system,

21702

it is assumed to be in equilibrium, i.e., the downward pull of gravity $m\mathbf{g}$ on the particle is balanced by the tension $\mathbf{F}_t = m\mathbf{g}$ with which the spring pulls upward on the mass. Now experiment shows that for small displacements from equilibrium (in general "small" will mean small compared with the length of the spring, though actually in this case the definition depends on the physical characteristics of the spring, a topic to which we shall recur later), the additional force \mathbf{F}_d (generally in the form of a weight) required to stretch the spring in the direction of its length and thus displace the mass particle is approximately directly proportional to the displacement. This is, of course, an illustration of Hooke's law. From the third law of Newton the same statement is true for the additional tension with which the spring pulls back on the mass, the so-called elastic restoring force. We should therefore expect that when the displacing force is suddenly removed, leaving the mass particle subject to an unbalanced restoring force varying directly as the distance from the equilibrium point, the resulting motion of the particle will be approximately simple harmonic about this point as center. This is indeed verified by experiment and the resulting oscillations are called the free oscillations of the system consisting of the spring and attached mass particle. Their frequency, the spring being assumed to be massless, is

$$\nu = \frac{1}{2\pi} \sqrt{\frac{K}{m}}, \qquad (2.2-13)$$

where K is the coefficient of proportionality between the displacing force and resultant displacement, i.e.,

$$F_d = Kx, (2.2-14)$$

and the equation of motion is then

$$m\ddot{x} = -F_d = -Kx. \tag{2-2-15}$$

This has the same form as $(2\cdot2-1)$ with c=-K. The quantity K is often known as the *stiffness* of the spring. The formula $(2\cdot2-13)$ for the free oscillation frequency is worthy of special comment because we shall meet it again in our study of vibration problems. The significant thing to be noted is the presence of the elastic or stiffness factor in the numerator of the radical and the mass in the denominator. Thus, for given mass the stiffer the

spring, the greater the frequency; on the other hand, for given stiffness, the larger the mass, the smaller the frequency.

In our treatment of the above illustration, we have neglected one important point. The theory implies that the simple harmonic motion of the mass particle attached to the spring never stops when once started. Experience, of course, rejects this result, for all such oscillations die down eventually when the system is left completely to itself. This indicates that our theory is of but approximate application and that to render it more complete we must assume a frictional or damping force acting on the mass in addition to the elastic restoring force. The solution of this problem will be postponed to Sec. 9·2.

It still remains our task to notice the case where the force though proportional to the displacement is always in the same direction as the displacement, that is, the constant c in eq. (2·2-1) is positive. The first integration then yields by the method used at the beginning of this section,

$$\frac{1}{2}m\dot{x}^2 = \frac{1}{2}cx^2 + C_1, \qquad (2.2-16)$$

whence we have

$$v = \dot{x} = \pm \sqrt{\frac{c}{m}} \sqrt{x^2 + \frac{2C_1}{c}}$$
 (2.2-17)

A second integration then gives

$$\log\left(x+\sqrt{x^2+\frac{2C_1}{c}}\right)=\pm\sqrt{\frac{c}{m}}\cdot t+C_2,$$

from which we may express x in explicit form as

$$x = \frac{1}{2}e^{C_2}e^{\pm\sqrt{c/m}t} - \frac{C_1}{c}e^{-C_2}e^{\mp\sqrt{c/m}t},$$

which may in turn be written in the more convenient form

$$x = Ae^{\sqrt{c/m}t} + Be^{-\sqrt{c/m}t},$$
 (2.2–18)

where the arbitrary constants A and B now replace C_1 and C_2 . We shall recall without proof from the study of differential equations that the general solution of a second order linear differential equation like $(2\cdot 2-1)$ contains two arbitrary constants. Indeed $(2\cdot 2-18)$ satisfies this requirement and the original differential equation, for c positive. It remains only to express A and B in

44

terms of the initial velocity and displacement of the particle. Taking the simple case in which x = R and v = 0 when t = 0, we have the conditions

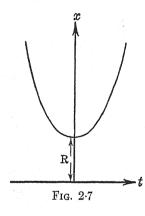
$$(1) A + B = R,$$

(2)
$$A - B = 0$$
,

whence A = B = R/2 and therefore

$$x = \frac{R}{2} \cdot [e^{\sqrt{c/m}t} + e^{-\sqrt{c/m}t}].$$
 (2-2-19)

This may be more concisely expressed in terms of hyperbolic functions, i.e., recalling that $e^x + e^{-x} = 2 \cosh x$, we may write



$$x = R \cosh \sqrt{\frac{c}{m}} t. \quad (2\cdot 2-20)$$

The displacement increases rapidly with the time and there is no oscillation. The graph of $(2\cdot 2-20)$ is shown in Fig. 2.7. The curve is known as the "catenary," and we shall have occasion to use it again in connection with problems in statics.

The type of force leading to $(2\cdot 2-20)$ is a repulsion from a fixed point. repulsive forces are encountered in phys-

ical phenomena, the one here considered is uncommon, and it will not be profitable to discuss it further. It has served indeed the useful purpose of stressing the method of handling initial conditions.

2.3. Motion in a Field Proportional to the Inverse Square of The equation of motion for a particle moving in a the Distance. straight line and acted on by a force inversely proportional to the square of the distance from a given point of the line, taken as origin, is

$$m\ddot{x} = \frac{c}{x^2} \cdot \tag{2-3-1}$$

If c is positive, the force is repulsive, while negative c corresponds to

an attractive force. We shall here discuss only the latter case, as it is physically by far the more important. Thus as in Sec. 2.2, we put c = -k, where k is a positive constant. The eq. $(2\cdot3-1)$ then takes the form, multiplying both sides by \dot{x} dt,

$$\frac{1}{2}md(\dot{x})^2 = kd\left(\frac{1}{x}\right), \qquad (2\cdot 3-2)$$

the result of the integration of which is

$$\frac{1}{2}m\dot{x}^2 = \frac{k}{x} + C_1, \qquad (2.3-3)$$

or solving for the velocity,

$$\dot{x} = \pm \sqrt{\frac{2k}{mx} + \frac{2C_1}{m}}$$
 (2.3-4)

Let us evaluate the constant C_1 by the boundary condition that the particle starts from rest at $x = R_0$. Then $C_1 = -k/R_0$, and eq. (2·3-4) becomes on separation of variables

$$\frac{dx}{\sqrt{1/x - 1/R_0}} = \pm \sqrt{\frac{2k}{m}} dt. \tag{2.3-5}$$

We must actually choose the minus sign before the radical on the right since positive dt will here correspond to negative dx. The result of a second integration may be expressed in the form

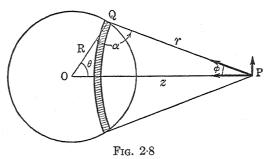
$$-\sqrt{R_0} \left(\sqrt{R_0 x - x^2} - R_0 \arcsin \sqrt{\frac{x}{R_0}} \right)$$

$$= -\sqrt{\frac{2k}{m}} \cdot t + C_2. \tag{2.3-6}$$

The boundary condition just assumed leads to $C_2 = \pi/2 \cdot R_0^{3/2}$. It follows that the time the particle takes to reach the origin if it starts from rest at $x = R_0$ is

$$t_1 = \sqrt{\frac{m}{2k}} \cdot \frac{\pi}{2} \cdot R_0^{3/2}. \tag{2.3-7}$$

According to Newton's law or hypothesis of universal gravitation every particle of matter in the universe attracts every other particle with a force, directed along the line joining the particles, directly proportional to the product of their masses and inversely proportional to the square of their distance apart, the coefficient of proportionality being known as the gravitation constant and generally denoted by G. The motion being discussed in this section is therefore linear motion in a gravitational field and of great importance in physics. In order to carry forward the discussion at this place, it will be necessary for us to interject a preliminary computation of the resultant attraction between an aggregate of mass particles



and another single particle. Let us consider the aggregate in the form of a homogeneous sphere and calculate its attraction for an external particle. The simplest procedure is to begin with a spherical shell with center at O (Fig. 2·8). The external mass particle, supposed for simplicity to be of unit mass, is located at P. Let OP = z and the radius of the shell be R, while its mass is m, assumed to be uniformly distributed over its surface. Now let us draw about P as center two spheres of radii r and r + dr. These will cut the shell in a surface ring element of area $2\pi R^2 \sin\theta \ d\theta$. For the width of the ring is $R \ d\theta$, where $d\theta$ is the angle between the radii to the two edges of the ring. The circumference of the ring is $2\pi R \sin\theta$. The area is the product. All points of the ring are at approximately the same distance from P.

We may now conceive this zonal surface element to be subdivided into small sub-elements of unit area (e.g., the one at Q), each of mass $m/4\pi R^2$ and equidistant from P. The force of attraction between each sub-element and the unit mass at P has then the magnitude

$$\frac{m}{4\pi R^2} \cdot \frac{G}{r^2},\tag{2.3-8}$$

and is directed from P to the sub-element in question. To find the resultant of these forces (i.e., the force between unit mass at P and the whole zonal surface element) we may most conveniently get their components normal and parallel to PO respectively. Examination at once discloses that the sum of the normal components is zero. Hence the required resultant is directed along PO. Its magnitude is obtained by multiplying the magnitude of each parallel component by the number of sub-elements. If the angle QPO is denoted by ϕ , the magnitude of each parallel component is

$$\frac{m}{4\pi R^2} \cdot \frac{G}{r^2} \cdot \cos \phi, \qquad (2.3-9)$$

while the number of sub-elements is simply the area $2\pi R^2 \sin \theta \, d\theta$. Hence the magnitude of the resultant attraction is

$$\frac{Gm}{2r^2} \cdot \sin \theta \cos \phi \, d\theta. \tag{2.3-10}$$

In order to find the complete attraction for the whole shell, this must be integrated as θ goes from 0 to π . However, θ does not appear to be a desirable integration variable, hence we shall use r. From the law of cosines

$$r^2 = R^2 + z^2 - 2Rz\cos\theta,$$

so that

$$\sin\theta \, d\theta = \frac{r \, dr}{Rz}.$$

Moreover

$$\cos\phi = \frac{z^2 - R^2 + r^2}{2zr}.$$

Therefore on substitution the expression in $(2\cdot3-10)$ takes the form

$$\frac{Gm}{4Rz^2} \left[1 + \frac{z^2 - R^2}{r^2} \right] dr. \tag{2.3-11}$$

This must be integrated between the limits r=z-R and r=z+R. The total attraction of the spherical shell for the unit

mass at P then has the magnitude

$$F = \frac{Gm}{4Rz^{2}} \left[r - \frac{z^{2} - R^{2}}{r} \right]_{z-R}^{z+R}$$

$$= \frac{Gm}{z^{2}} \cdot \tag{2.3-12}$$

This means that the spherical shell attracts the external point as if all its mass were concentrated at its center. Now any spherical mass may be thought of as made up of a series of concentric spherical shells of masses m_1, m_2, m_3, \ldots From what we have just proved, the attraction for an external point of the totality of such shells has the magnitude

$$\frac{G(m_1 + m_2 + m_3 \cdots)}{z^2} = \frac{GM}{z^2}, \qquad (2-3-13)$$

where M is the total mass of the sphere. It therefore follows that a solid sphere like a spherical shell attracts an external mass particle as if all its mass were concentrated at its center. For example, this will be very nearly true of the earth which is approximately spherical.

We are now in a position to make more specific application of the earlier results of this section. Let us suppose that a particle of mass m falls to the earth from rest at a height h above the surface. The velocity which it has obtained at the moment of striking, neglecting air resistance, is given by eq. (2·3-4), substituting x = R and k = GmM, where M is the mass of the earth. For the constant C_1 , we substitute -GmM/(h+R), where R is the radius of the earth and it is assumed that the velocity is zero at x = h + R. Hence the velocity at the moment of impact is

$$v_0 = \sqrt{\frac{2GMh}{(Rh + R^2)}}. (2.3-14)$$

In particular let us further suppose that $h \gg R$ so that R^2/h can be neglected in comparison with R. The velocity in question, which we may call the velocity from infinity, then takes the simple form

$$v_{\infty} = \sqrt{\frac{2GM}{R}} \cdot \tag{2.3-15}$$

The same quantity represents the velocity with which a mass particle must be ejected from the earth's surface never to return.

Its magnitude with $G=6.67\times 10^{-8} \frac{\mathrm{dynes}}{\mathrm{gm^2}} \mathrm{cm^2},~M=6.1\times 10^{27}$

gms, and $R=6.37\times 10^8$ cm turns out to be 1.13×10^6 cm/sec or approximately 7 miles/sec.

It is interesting to observe that if $h \ll R$, (2·3-14) now yields

$$v_0 = \sqrt{\frac{2GMh}{R^2}} \cdot \tag{2.3-16}$$

But near the surface of the earth

$$\frac{GmM}{R^2} = mg, (2.3-17)$$

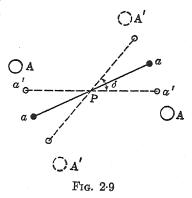
where g is the acceleration of gravity at points near the surface. Hence v_0 becomes

$$v_0 = \sqrt{2gh}, \qquad (2.3-18)$$

which is the usual formula for the velocity attained by a particle falling freely from a height h in the neighborhood of the earth's

surface. This will enable the student to connect the present discussion with the simpler case developed in Sec. 2-1.

The gravitation constant "G" which has entered so prominently into the discussion of the present section is one of the most important constants of physics. It is therefore of interest to note the most sensitive method for the determination of this quantity. This employs the so-called torsion balance in which two small metal



spheres (usually of platinum or gold) are fixed to the ends of a light rod. This in turn is suspended at its mid-point by a fine vertical quartz fiber. In the schematic diagram (Fig. 2.9) aPa denotes the rod with the two metal balls at its extremities in its equilibrium position. P is the point of suspension by the fiber which is, of course, perpendicular to the plane of the diagram. Two heavy

lead spherical balls are now placed at A, A and as a result the rod is deflected to a'Pa' as indicated by the dotted lines. When the lead balls are moved to A', A', the deflection is in the opposite direction and the total deflection is measured by the angle δ . From a knowledge of the elastic properties of the fiber (see Sec. 10·1) the magnitude of the force of attraction F between the small

and large spheres can be computed and G calculated from the equation

 $F = \frac{GmM}{r^2},$

where m and M are the masses of the small and large spheres respectively and r is the distance between their centers. This method was devised in the late eighteenth century by Henry Cavendish and his name is usually attached to the experiment. The most recent exact measurements by this method are due to P. R. Heyl of the United States Bureau of Standards.1 Hisapparatusis shown in the adjacent figure taken from

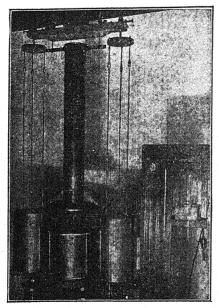


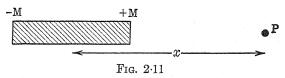
Fig. 2.10

the article indicated in the footnote (Fig. 2·10). Heyl used small platinum spheres of mass about 50 grams each and large steel cylinders of mass 66 kilograms each. For greater accuracy the torsion pendulum was placed in an evacuated region in which the pressure was usually about 1 or 2 mm. of Hg. Heyl allowed the rod to swing as a torsion pendulum instead of following the earlier static deflection method. His value of G, which is now

the accepted one, is $(6.670 \pm 0.005) \times 10^{-8} \frac{\text{dyne cm}^2}{\text{gm}^2}$.

¹ P. R. Heyl, Proceedings of the National Academy of Sciences, Volume 13, p. 601 (1927).

2.4. Motion in a Field Proportional to the Inverse Cube of the Distance. This is a case of some physical interest because it arises when we have two force centers, very close together, the one attracting and the other repelling a distant particle with a force proportional to the inverse square of the distance. An example from magnetism will sufficiently illustrate this. Imagine a very short magnet of length l and with north pole +M and south pole



-M (Fig. 2·11), and inquire the resultant force on a unit north pole placed at P, distant x from the center of the magnet along the magnet extended, where $x \gg l$. From Coulomb's law, according to which the force which a single isolated magnetic pole of strength M poles exerts on another of strength M' at a distance x is MM'/x^2 , we have for the magnitude of the resultant force on the unit pole at P (in this case a force of repulsion)

$$\frac{-M}{(x+l/2)^2} + \frac{M}{(x-l/2)^2} = \frac{M \cdot 2xl}{(x^2-l^2/4)^2}, \qquad (2.4-1)$$

where the first term on the left denotes the force due to the south pole and the second that due to the north pole, and on the right we have reduced to a common denominator. Now we further have

$$\frac{M\cdot 2xl}{(x^2-l^2/4)^2} \doteq \frac{2Ml}{x^3},$$

if we utilize the assumption that $l \ll x$. The direction of the force is away from and in the line containing the magnet.

In a force field of this nature, if the moving particle has mass m, the equation of rectilinear motion is

$$m\ddot{x} = \frac{K}{x^3}, \qquad (2.4-2)$$

where K is a positive constant (e.g., 2Ml in the above illustration). Conducting a first integration, we have



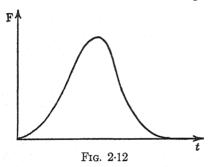
$$\frac{1}{2}m\dot{x}^2 = \frac{-K}{2x^2} + C_1. \tag{2.4-3}$$

Suppose that the particle starts from rest at a distance R from the force center. Then $C_1 = K/2R^2$. The velocity which the particle will have after traveling to infinity will then be

$$v = \frac{1}{R} \sqrt{\frac{K}{m}}, \qquad (2.4-4)$$

or inversely proportional to the distance from which it started. The student should also investigate the corresponding case in which K is a negative constant. (Problem 15 at the end of the chapter.)

2.5. Force Dependent on the Time. Impulsive Force. The forces considered so far in this chapter have been either constant or



dependent on the distance from some chosen origin. These are not the only forces leading to motions of interest in physics. Let us recall the type of force involved in the blow of a hammer. Here the force is definitely a function of time, rising from zero to a maximum and again falling to

zero, more or less as in the accompanying figure (Fig. 2·12). If the time interval during which the force acts is sufficiently small we refer to the force as *impulsive*. Forces associated with the collisions of particles are usually of this character. The shape of the force curve is not always known, but a typical one might be of the form

$$F = Ate^{-Bt} (2.5-1)$$

This is something like Fig. 2·12, except that F does not vanish for finite t but goes asymptotically to zero as t increases. If B is sufficiently large, however, F approaches zero fast after attaining its maximum, which occurs for t=1/B. The equation of motion of a particle of mass m in this case is

$$m\ddot{x} = Ate^{-Bt}. (2.5-2)$$

A first integration yields the velocity, viz.:

$$v = \dot{x} = \frac{A}{m} e^{-Bt} \left(-\frac{t}{B} - \frac{1}{B^2} \right) + C,$$
 (2.5-3)

where C is the constant of integration. In terms of the initial velocity v_0 this becomes

$$v - v_0 = \frac{A}{m} \left[-\frac{t}{B} e^{-Bt} + \frac{1}{B^2} (1 - e^{-Bt}) \right].$$
 (2.5-4)

According to this, v attains a limiting velocity $A/B^2m + v_0$ which is non-vanishing, though it can be very small if $v_0 = 0$ and B is small enough. The maximum velocity is obtained in the usual way by setting dv/dt = 0. It then turns out that this happens only for $t \to \infty$, and hence the limiting velocity is the maximum velocity. The determination of the distance traveled in time t is left as an exercise for the reader.

The action of most impulsive forces encountered in practice is more complicated than that considered above, the reason being that the motion produced by the impulsive force is commonly resisted by another force whose value changes during the motion and is generally dependent on the velocity. This leads to the considerations of the next section.

2.6. Rectilinear Motion in a Resisting Medium. It is found by experience that when a force produces acceleration through an actual material medium like the air (e.g., motion of a rain drop) or water (e.g., motion of a ship) or even a solid (as in the motion of a nail through a piece of wood) the medium always resists the motion with a force varying as some power of the velocity. The simplest case is that in which the resisting force varies directly as the velocity but is oppositely directed. This turns out to be a fair representation of the facts for motion in still air, for example, if the average velocity is small.

Let us consider the simple problem of a ball of mass m tossed straight up in the air. With the x axis vertical we have for the equation of motion

$$m\ddot{x} = -mg - R\dot{x}, \qquad (2.6-1)$$

where R is a positive constant. Dividing through by m and

putting R/m = D for simplicity, we get the simpler form

$$\ddot{x} + D\dot{x} + g = 0. (2.6-2)$$

We let $\dot{x} = v = \text{instantaneous velocity}$, and have

$$\dot{v} + Dv + g = 0. {(2.6-3)}$$

We separate the variables by writing this as follows:

$$\frac{dv}{v+g/D} = -D dt, \qquad (2.6-4)$$

which has the solution

$$\log (v + g/D) = -Dt + K, (2.6-5)$$

where K is the constant of integration. Let the initial velocity be v_0 . Then $(2\cdot6-5)$ yields for the instantaneous velocity

$$v = (v_0 + g/D)e^{-Dt} - g/D.$$
 (2.6-6)

It is of interest to see what this becomes for D very small, so that we can expand e^{-Dt} in the usual series

$$e^{-Dt} = 1 - Dt + D^2t^2/2! - D^3t^3/3! + \cdots$$
 (2.6-7)

and keep only the first three terms. The result is

$$v = v_0 - gt - v_0Dt + (v_0 + g/D)D^2t^2/2! + \cdots$$
 (2.6-8)

As $D \rightarrow 0$, this takes on the familiar form

$$v = v_0 - gt \tag{2.6-9}$$

with the time of rising into the air equal to v_0/g . In the more general case in which D is non-vanishing the time to reach the highest point h is

$$t_h = \frac{1}{D} \log (1 + v_0 D/g).$$
 (2.6-10)

To compare this with v_0/g in the non-resistance case, we again expand in a series (treating D as small) and obtain

$$t_h = v_0/g - v_0^2 D/2g^2 + \cdots$$
 (2.6-11)

It thus appears that $t_h < v_0/g$ or the time to reach the highest point is less in the resisting medium than in vacuo. This should be more carefully verified by the use of the inequality

$$\log (1+x) < x,$$

which is true for all real, positive values of x. The result appears a trifle surprising until we reflect that the total height of rise will probably turn out to be less than in the non-resistance case. To get this we must solve $(2\cdot6-6)$ by putting $v=\dot{x}$ again, thus obtaining

$$dx = [(v_0 + g/D)e^{-Dt} - g/D] dt. (2.6-12)$$

The integration yields (using the initial condition x = 0 for t = 0)

$$x = \left(\frac{v_0 + g/D}{D}\right) (1 - e^{-Dt}) - \frac{gt}{D}.$$
 (2.6-13)

Once more, expansion in series gives the approximate result

$$x = v_0 t - g t^2 / 2 - v_0 D t^2 / 2 + \cdots$$
 (2.6-14)

As $D \to 0$ this takes on the familiar form for upward flight in vacuo. We obtain the maximum height h by substituting t_h for t in (2·6–13). The result is

$$h = v_0/D - g/D^2 \cdot \log (1 + v_0 D/g).$$
 (2.6-15)

Expansion of the logarithm again puts this in a form more readily compared with the ordinary ideal case of motion in vacuo. Thus

$$h = v_0^2/2g - v_0^3D/3g^2 + \cdots$$
 (2.6-16)

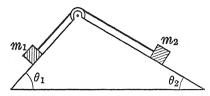
The second term on the right thus appears as a correction (for D small) to the ideal maximum height given by the first term. The indication is thus that the maximum height attained in a resisting medium for given initial velocity v_0 is less than in the non-resisting medium. This is verified by more careful examination of (2.6-15).

The use of series expansions in this section is recommended to the reader as a useful device in many mechanical problems.

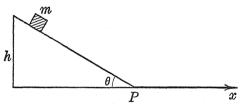
PROBLEMS

- 1. A monkey climbs a massless rope which passes over a massless and frictionless pulley P and has fastened on the other end a counter weight W of the same mass as the monkey. Discuss the motion of the monkey and the weight.
- 2. In the accompanying figure particles of mass m_1 and m_2 respectively are connected by a perfectly flexible inextensible string and are constrained to move on the two inclined planes as indicated. Obtain the expression for the acceleration of the system, as well as that for the tension in the string. If the

system starts from rest, what is the kinetic energy after each particle has moved a distance s along its respective plane?



3. In the figure a particle of mass m starting from rest slides down a perfectly smooth inclined plane of angle θ and height h. When it reaches P, it continues to move along the horizontal in the x direction but subject to a resisting force of the form kv, where v is the speed and k is a constant. How far will it go before coming to rest? How long will it take to cover 99% of this distance?



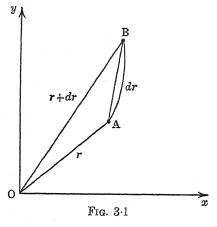
- 4. Prove that the time for a particle to descend a smooth inclined plane is to the time to descend vertically a distance equal to the height of the plane as $1 : \sin \theta$, where θ is the angle of the plane.
- 5. A metal ball of mass 100 grams falls from rest through a column of glycerine. It is ultimately observed to attain a practically constant velocity of 5 cm/sec. If the resisting force of the glycerine is assumed to vary as the first power of the velocity of the ball, find the magnitude of the force for any velocity v. Also find the distance traveled in the first second.
- 6. A particle of mass M is subject to two forces: (1) a constant force F; and (2) a resisting force varying as the square of the velocity. Find the expression for the distance traveled by the particle in time t and its velocity at the same time.
- 7. Prove that the motion of the projection of a fixed point located on the circumference of a circle on a uniformly rotating diameter of the circle is, relative to the diameter, simple harmonic motion.
- 8. A straight smooth tube is bored through the center of an otherwise homogeneous solid sphere of mass M and radius R. A particle of mass m falls from rest through the tube from the surface of the sphere. What will its motion be, assuming that the only force acting is the gravitational force between the particle and the sphere? How long will it take the particle to travel through the sphere?

- 9. One end of a vertical brass wire 10 meters long and 1 mm in radius is fastened to a rigid ceiling. The other end is loaded with a mass of 2 kilograms. If the mass is displaced 5 mm and released, show that its resulting motion is simple harmonic. If Young's modulus for brass is 9×10^{11} dynes/cm², find the frequency and the maximum kinetic energy of the load. What is the impulse of the force during a quarter period of the motion? How much average power is exerted during a quarter period?
- 10. A cylindrical disc of radius R and height h floats in water with half its volume immersed, and its flat side horizontal. If it is depressed a distance $x \ll h$ and released, show that the resulting motion is simple harmonic and find the frequency.
- 11. A body falls 2000 miles to the surface of the earth. Find the time it takes and also the time for the last 1000 miles.
- 12. An electron is attracted by a positive charge of magnitude Ne. Suppose that in addition to this there also acts a repulsive force on the electron varying inversely as the cube of the distance. Determine the coefficient of proportionality of this latter force so that the resultant force on the electron is zero at a distance of 10^{-11} cm from the positive charge. What velocity will it have at this point if it starts at rest from a distance of 10^{-8} cm?
 - 13. Plot a space-time graph for the motion of the electron in Problem 12.
- 14. A free particle (i.e., one which is subject to no accelerating forces) possesses initial velocity v_0 in a straight line in a medium which exerts a resisting force proportional to the cube of the velocity. Show that the time it takes the particle to travel a given distance is a quadratic function of the distance.
- 15. A particle of mass m moves in a straight line subject to a force directed toward a fixed point on the line and varying inversely as the cube of the distance. Discuss the motion. In particular if the particle starts from rest at distance h from the fixed point, find the total time required to reach the latter point.
- 16. A particle of mass m is acted on by an attractive force directed along the x axis and varying inversely as the nth power of the distance from the origin, where n is integral and different from 0 or 1. If the particle starts at rest at distance a from the origin, find the velocity it will attain at distance x.

CHAPTER III

CURVILINEAR MOTION IN A PLANE

3.1. Components of Motion in a Plane. After rectilinear motion the next stage of complexity is clearly motion in a plane. It follows from the discussion in Sec. 1.2 that the position of a particle in a plane may be represented either by the position vector \mathbf{r} from



some chosen origin to the point occupied by the particle or by the rectangular coördinates x, y in an appropriate reference system. In either case we have the relation

$$\mathbf{r} = \mathbf{i}x + \mathbf{j}y, \quad (3\cdot 1-1)$$

where i and j are the *unit* vectors along the x and y axes respectively.¹ To consult the figure (Fig. 3·1), if AB denotes an infinitesi-

mally small portion of the path of a particle, the vector from O to A which gives the position of A will be denoted by \mathbf{r} and the corresponding vector to B is $\mathbf{r} + d\mathbf{r}$. Hence the displacement AB may be represented by the vector $d\mathbf{r}$. The velocity of the particle is $\mathbf{v} = \dot{\mathbf{r}}$ (Sec. 1-3), which from (3-1-1) becomes

$$\mathbf{v} = \dot{\mathbf{r}} = \dot{\mathbf{i}}\dot{x} + \dot{\mathbf{j}}\dot{y}. \tag{3.1-2}$$

It should be emphasized that in differentiating r in (3·1-1) with respect to the time the unit vectors i and j are treated as constants since the axes are assumed to remain fixed as time passes.

Similarly the acceleration of the particle is by definition

$$\mathbf{a} = \dot{\mathbf{v}} = \ddot{\mathbf{r}} = \mathbf{i}\ddot{x} + \mathbf{j}\ddot{y}, \tag{3.1-3}$$

 $^{^{1}}$ Eq. (3·1-1) is thus a special case of (1·3-6).

whence \ddot{x} and \ddot{y} appear as the rectangular components of the vector acceleration. We may therefore write (cf. Sec. 1.5)

$$a_x = \ddot{x},$$

$$a_y = \ddot{y}.$$
(3·1-4)

We shall have occasion to use these acceleration components a great deal in the following sections.

3.2. Equations of Motion for a Particle in a Plane. The motion of a particle in a plane is a special case of motion in space and hence is described by the fundamental equation (1.7-4). It is important to recall that this equation, viz.,

$$\mathbf{F} = m\ddot{\mathbf{r}},\tag{3.2-1}$$

is a vector equation. Its solution for each particular choice of ${\bf F}$ involves the integration of a second order ordinary differential equation. Methods are available in certain cases for the integration in vector form, but we shall prefer the more common alternative method of components. Thus we write $(3\cdot2-1)$ in the form (specialized to the xy plane)

$$iF_x + jF_y = m(ia_x + ja_y), \qquad (3.2-2)$$

where F_x and F_y are the x and y component forces already mentioned, and a_x and a_y similarly the component accelerations. Since (3·2-2) holds in general, the coefficient of i on the left must be equal to that of i on the right, etc., with the result that (3·2-2) is actually equivalent to the two equations

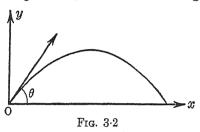
$$F_x = ma_x = m\ddot{x},$$

$$F_y = ma_y = m\ddot{y}.$$
 (3.2-3)

These are called the component equations of motion in the plane. Their independent solution provides the solution of the problem of plane motion. It is well to stress that our handling of them as independent simultaneous equations reflects our faith that a force F_x can act on the particle in the x direction independently of the action of F_y in the y direction. This is, of course, an illustration of the principle of superposition and is implicit in the vector form of the fundamental equation of motion.

We now proceed to a simple illustration of the integration of equations of the form $(3\cdot2-3)$.

3.3. Projectile Motion. Consider a particle of mass m projected into the air from a point on the ground with velocity of magnitude v_0 in a direction making an angle θ with the horizontal.



We wish to determine its motion, neglecting the resistance of the air, which for small velocities will not affect the motion appreciably. We also neglect any effect of the wind. Since there are then no forces save gravity, the motion is confined to a

plane which we shall take as the xy plane (cf. Fig. 3.2). If we draw the x and y axes through O, the point of projection, the force components are then

$$F_x = 0, \quad F_y = -mg,$$
 (3.3-1)

where we assume that the motion takes place so close to the surface of the earth that we are justified in taking the acceleration g as a constant. To be perfectly general we should have to take account of the fact that the particle is attracted to the earth by the Newtonian force of gravitation and hence should use the inverse square law. However, from the discussion in Sec. 2-3 it is clear that for elevations small compared with the radius of the earth, the error in assuming a constant g in $(3\cdot3-1)$ is negligible compared with that made in neglecting air resistance.

The resulting equations of motion are (cf. 3.2-3)

$$m\ddot{x} = 0,$$

$$m\ddot{y} = -mg.$$
(3·3-2)

Integrating, we have at once

$$x = \frac{c_1}{m}t + c_2, (3.3-3)$$

$$y = -\frac{1}{2}gt^2 + c_3t + c_4, (3.3-4)$$

where c_1 , c_2 , c_3 and c_4 are the usual arbitrary constants of integra-

tion. To evaluate them, let us note that at t = 0, x = y = 0, and therefore $c_2 = c_4 = 0$. Moreover

$$\frac{c_1}{m} = \dot{x} \bigg|_{t=0} = v_0 \cos \theta, \tag{3.3-5}$$

and

$$c_3 = \dot{y}|_{t=0} = v_0 \sin \theta. \tag{3.3-6}$$

The component velocities at any instant are then

$$\dot{x} = v_0 \cos \theta, \quad \dot{y} = -gt + v_0 \sin \theta, \tag{3.3-7}$$

and the final parametric equations of the path of the particle are

$$x = v_0 \cos \theta \cdot t, \quad y = -\frac{1}{2}gt^2 + v_0 \sin \theta \cdot t.$$
 (3.3-8)

Eliminating t between the two equations in (3.3-8) we get the equation of the path of the projectile in the form

$$y = -\frac{1}{2}g \frac{x^2}{v_0^2 \cos^2 \theta} + x \tan \theta,$$
 (3.3-9)

which is the equation of a parabola with axis parallel to the y axis and with vertex at the point with the coördinates

$$x = \frac{v_0^2 \sin 2\theta}{2g}$$
, $y = \frac{v_0^2 \sin^2 \theta}{2g}$.

This is the point at which the particle reaches its maximum height, as may be verified by applying the usual test of equating dy/dx to zero. The maximum height is therefore

$$h = \frac{v_0^2 \sin^2 \theta}{2q}, (3.3-10)$$

and the range, or total distance which the projectile travels along the x axis, is, from symmetry,

$$R = \frac{v_0^2 \sin 2\theta}{g} \cdot \tag{3.3-11}$$

For a given value of v_0 , R is a maximum for $\theta = 45^{\circ}$. The symmetry of the parabola about its axis indicates that the time spent on the upward part of the flight is equal to the time for the down-

ward flight. From (3·3-8) the total time spent by the projectile in its flight is readily seen to be

$$t_{\mathbb{R}} = \frac{2v_0 \sin \theta}{g} \cdot \tag{3.3-12}$$

In our discussion of the motion of a particle under gravity both in this section and Sec. 2·1, we have deliberately overlooked one point of small practical but considerable theoretical significance. It will recalled from Sec. 1·8 that the equations of motion which we have used are applicable only in inertial systems, i.e., those which move with constant velocity (without rotation) with respect to the primary inertial system. In treating axes fixed in the rotating earth as an inertial system we are therefore committing a logical error and our results cannot then be completely correct. The thorough solution of the problem will be found in Sec. 7·12, where the correct equations of motion for such a case are set up and solved. As has been intimated, the modification in the results of the present section will be found to be very slight in practice.

3.4. Projectile Motion in a Resisting Medium. The general problem of the plane motion of a projectile through a medium that resists its flight is a highly involved one which it is the function of exterior ballistics to solve. Nevertheless there is some value in considering here the rather simple special case in which the resisting force varies as the square of the resultant velocity magnitude and is directed at every point along the trajectory oppositely to the direction of motion.

Let us assume that the magnitude of the resisting force is of the form

$$kmv^2$$
,

where m is the mass of the particle, and k is the specific resistance factor, i.e. the resisting force per unit mass per unit (velocity)². The component equations of motion then become

$$\ddot{x} = -k\dot{x}\dot{s},$$

 $\ddot{y} = -k\dot{y}\dot{s} - g,$ (3·4-1)

where

$$\dot{s}^2 = \dot{x}^2 + \dot{y}^2 \tag{3.4-2}$$

63

is the square of the magnitude of the resultant velocity. We find it convenient to use as a variable, s, which denotes distance measured from the origin along the path of the projectile. The reader should be able to show without difficulty from the first equation in $(3\cdot 4-1)$ that

$$x = v_0 \cos \theta \cdot e^{-ks}, \qquad (3.4-3)$$

where v_0 and θ have the same meaning as in Sec. 3.3. Note that when s=0 (i.e. at the origin or point of projection) $\dot{x}=v_0\cos\theta$, the initial horizontal velocity. Combination of (3.4–3) with the second equation in (3.4–1) gives the differential equation of the orbit in the form

$$\frac{d^2y}{dx^2} = -\frac{g}{v_0^2 \cos^2 \theta} \cdot e^{2ks}.$$
 (3.4-4)

Note that for k=0, i.e. no resistance, this reduces to a differential equation which when integrated yields precisely the parabolic orbit $(3\cdot3-9)$ as it should.

If we now restrict the discussion to a rather flat trajectory, i.e. one in which $|dy/dx| \ll 1$ in all parts of the path, the first integral of (3·4-4) becomes

$$\frac{dy}{dx} = \tan\theta + \frac{g}{2kv_0^2\cos^2\theta} \cdot (1 - e^{2kx}), \qquad (3.4-5)$$

yielding on a second integration the equation of the orbit in the form

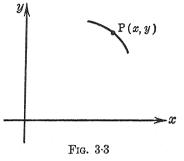
$$y = x \left(\tan \theta + \frac{g}{2kv_0^2 \cos^2 \theta} \right) + \frac{g}{4k^2v_0^2 \cos^2 \theta} \cdot (1 - e^{2kx}) \quad (3.4-6)$$

where the constant of integration has already been evaluated by the condition that, when x = 0, y = 0. If we expand e^{2kx} into the usual series and assume that kx is small so that no terms need be kept after the third power ones, the orbit becomes

$$y = x \tan \theta - \frac{gx^2}{2v_0^2 \cos^2 \theta} - \frac{gkx^3}{3v_0^2 \cos^2 \theta}.$$
 (3.4-7)

The last term on the right represents the correction to the ideal parabola due to the presence of resistance. This might serve as an approximation to the path of a rifle bullet fired through air with a small elevation angle θ . Ballistically speaking the result should not be taken too seriously, however, since the resisting force does not vary with the square of the velocity for all speeds, and is not

directed exactly opposite to the velocity of the bullet.



3.5. Composition of Simple Harmonic Motions in a Plane. Certain types of motion of interest in physics arise from the composition of simple harmonic motions along the coördinate axes. Confining our discussion as usual to motion in a plane, let

us suppose that a mass particle at P (Fig. 3-3) moves in accordance with the following equations of motion

$$m\ddot{x} = -c_1 x,$$

$$m\ddot{y} = -c_2 y,$$

$$(3.5-1)$$

where c_1 and c_2 are positive constants. Consulting Sec. 2.2 we see that each of these can be integrated directly, yielding

$$x = A_1 \cos \left(\sqrt{\frac{c_1}{m}} t + \epsilon_1 \right),$$

$$y = A_2 \cos \left(\sqrt{\frac{c_2}{m}} t + \epsilon_2 \right).$$
(3.5-2)

There are thus two simple harmonic motions along the x and y axes respectively having amplitudes A_1 and A_2 and frequencies

$$v_1 = \frac{1}{2\pi} \sqrt{\frac{c_1}{m}} \text{ and } v_2 = \frac{1}{2\pi} \sqrt{\frac{c_2}{m}},$$

and differing in initial phase by $\epsilon_1 - \epsilon_2$. The equations (3.5–2) are the parametric equations of the path of the particle. We may obtain the geometric path by the elimination of the time between them.

Let us take first the special case where $c_1 = c_2 = c$, i.e., where the frequencies are the same along the two axes. Let $\sqrt{c/m} = \omega$ and for convenience choose the initial phase ϵ_1 to be zero by taking

the initial x displacement equal to A_1 . Then the total phase difference between the two motions will be ϵ_2 , which we shall call ϵ for simplicity. The parametric equations become

$$x = A_1 \cos \omega t, y = A_2 \cos (\omega t + \epsilon)$$
 (3.5–3)

By substitution from the first into the second, t is eliminated and we have

$$y = A_2 \left[\frac{x \cos \epsilon}{A_1} - \sqrt{1 - \frac{x^2}{A_1^2}} \cdot \sin \epsilon \right].$$

By transposition, squaring and collecting terms, this equation takes the form

$$\frac{x^2}{A_1^2 \sin^2 \epsilon} - \frac{2xy}{A_1 A_2 \sin \epsilon \tan \epsilon} + \frac{y^2}{A_2^2 \sin^2 \epsilon} = 1. \quad (3.5-4)$$

Applying the usual test for a conic section, we find that this is the equation of an ellipse. Thus for an equation of the second degree in the form

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0,$$

we have a hyperbola, ellipse or parabola according as $B^2 - 4AC$ is greater than, less than or equal to zero respectively. In the present case this quantity is $-4/(A_1^2A_2^2\sin^2\epsilon)$ and is less than zero. The student should show that the major axis of the ellipse lies along the line making with the x axis the angle θ such that

$$\tan 2\theta = \left\{ \frac{2A_1 A_2 \cos \epsilon}{A_1^2 - A_2^2} \right\} \cdot \tag{3.5-5}$$

Let us note a few special cases. If $A_1 \neq A_2$ and $\epsilon = 90^{\circ}$, $\theta = 0$ and the major axis of the ellipse lies in the x axis. Supposing further $A_1 = A_2 = A$ and $\cos \epsilon = 0$, i.e., $\epsilon = \pi/2$, the ellipse reduces to a circle of radius A about the origin as center. The student should work out other cases and show for example that if

² See, for example, Smith and Gale, loc. cit., p. 182.

¹ See, for example, any text on analytic geometry, e.g., Smith and Gale, "New Analytic Geometry" (Ginn & Co., 1912), p. 179 ff.

 $\epsilon = 0$, the path is the straight line with equation

$$\frac{x}{A_1} = \frac{y}{A_2},\tag{3.5-6}$$

as is, of course, immediately evident from eq. (3.5-3).

It may be pointed out that the above motions are really special cases of a very important general kind of motion. For we see that if a particle is attracted toward the origin with a force proportional to the distance to the origin, i.e., if the force magnitude F is such that

$$F(r) = cr, (3.5-7)$$

where $r = \sqrt{x^2 + y^2}$, the component equations of motion will be precisely the equations (3·5-1), with $c_1 = c_2 = c$. Such a motion is a special case of what is called *central motion* or motion in a *central* field. It will be discussed in detail in the next section (Sec. 3·6), where the truth of the statement immediately above will become perhaps more evident.

Let us now, however, consider the case of eqs. (3.5-1) with $c_1 \neq c_2$. This will correspond to the composition of two simple harmonic motions with different frequencies. For simplicity we shall call these $\sqrt{c_1/m} = \omega$ and $\sqrt{c_2/m} = \omega + \alpha$, where α appears as the difference between the two angular frequencies. Now assume that there is a time when the two motions are in phase and choose this as the time origin. Then in eqs. (3.5-2) we have $\epsilon_1 = \epsilon_2 = 0$ and the equations become, choosing $A_1 = A_2 = A$ for greater simplicity,

$$x = A \cos \omega t,$$

$$y = A \cos (\omega + \alpha)t.$$
(3.5-8)

We first eliminate ωt and obtain

$$x^2 - 2xy \cos \alpha t + y^2 = A^2 \sin^2 \alpha t. \tag{3.5-9}$$

To get rid of the product term, change the coördinates to x', y', where

$$x = \frac{\sqrt{2}}{2} (x' - y'),$$

$$y = \frac{\sqrt{2}}{2} (x' + y').$$
(3.5-10)

This corresponds to a rotation of axes through the angle 45° . In the new coördinates the equation (3.5-9) becomes

$$x'^{2}(1 - \cos \alpha t) + y'^{2}(1 + \cos \alpha t) = A^{2}\sin^{2} \alpha t$$

or

$$\frac{x'^2}{2A^2\cos^2\alpha t/2} + \frac{y'^2}{2A^2\sin^2\alpha t/2} = 1, (3.5-11)$$

where we have used the trigonometric identities

1 -
$$\cos \alpha t = 2 \sin^2 \alpha t/2$$
; 1 + $\cos \alpha t = 2 \cos^2 \alpha t/2$; $\sin^2 \alpha t = 4 \sin^2 \alpha t/2 \cdot \cos^2 \alpha t/2$.

Equation (3.5–11) is the equation of an ellipse with axes which change continuously with the time and with values running from 0 to $2\sqrt{2}A$.

As a matter of fact, of course, eq. (3.5–11) cannot be used alone to determine the actual path which is the composition of the two simple harmonic motions. For this purpose the time must be eliminated between it and one of the equations (3.5–8). To plot the path in special cases it is most simple to use eqs. (3.5–8) as parametric equations of the orbit. As an illustration the case where $(\omega + \alpha)/\omega = 8/7$ or $\alpha = \omega/7$ is plotted in the accompanying figure (Fig. 3.4), the time origin being taken, of course, as that time when the motions are in phase.

Actually the figure does not plot precisely (3.5-8) but rather

$$x = -A \cos \omega t,$$

$$y = -A \cos (\omega + \alpha)t,$$

as is clear from the way the axes are chosen. This, however, is not significant.

A more important question is this: Is the resultant motion in this case periodic, i.e., does the path of the particle repeat itself after some period P? It is seen from (3.5-8) that if there is to be such a period P there must exist a pair of integers j and k such that

$$\omega P = 2\pi j,$$

$$(\omega + \alpha)P = 2\pi k.$$

¹ Taken from Leigh Page, "Introduction to Theoretical Physics" (D. Van Nostrand Co., N. Y., 1928), p. 72.

But this means that the ratio of the two frequencies ω and $\omega + \alpha$ must be the ratio of two whole numbers, i.e. $(\omega + \alpha)/\omega$ must be a rational number or put in another way, $\omega + \alpha$ and ω must be commensurable. This is certainly true in the example illustrated in Fig. 3.4. If $\omega + \alpha$ and ω are not commensurable, it follows that the resulting motion is not periodic.

Curves for both commensurable and incommensurable cases have been widely studied and are known as Lissajous figures. A

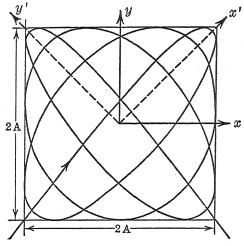


Fig. 3.4

simple method of obtaining them experimentally is to allow a stiff rod with rectangular cross-section to oscillate perpendicular to its length when clamped in a vise. If a mirror is attached to the freely vibrating end and a ray of light is reflected from it, the figures can then be produced on a screen. These figures are also obtained in the cathode ray oscillograph. In this instrument, the cathode rays are permitted to pass through two perpendicular oscillating electric or magnetic fields, and the resulting path traced by the rays as they pass through these two fields may be observed on a fluorescent screen at the other end of the tube.¹

The reader may find it of interest to express the horizontal and

¹ For an elaborate description of this instrument see L. B. Loeb, "Fundamentals of Electricity and Magnetism" (Wiley, New York, 1931), p. 281 ff.

CENTRAL FORCES

vertical harmonic motions in the form

$$x = A \cos \omega_1 t,$$

$$y = A \cos (\omega_2 t + \delta),$$



where the two angular frequencies ω_1 and ω_2 are different and the initial phase angle δ is not necessarily zero. As a matter of fact the Lissajous figure for $\delta=\pi/2$ is particularly useful since it turns out in this case that if ω_1 and ω_2 are commensurable, their ratio is equal to the ratio of the number of horizontal maxima to the number of vertical maxima in the figure. Electronic technicians make good use of this case.

3.6. Central Forces. Suppose that the force F on the particle is of such a nature that it is always directed toward or away from some fixed point, which we shall take as the origin, and is a function of the distance r from this point only. Such a force is termed a central force, and it is no exaggeration to state that central forces are among the most important of all that arise in physical problems. Geometrical reasoning shows that central field motion always takes place in a plane. Thus, the initial direction of motion of the particle and the initial line connecting it to the force center define a plane. Since there is no force component normal to this plane, motion once started in it will continue in it.

We shall call the line joining the particle P to the origin the radius vector. It is strictly \mathbf{r} , the position vector of the point P with respect to O. (Recall Sec. 1·2.) Its magnitude is r, which will always be considered a positive quantity. Consulting Fig. 3·5, let F(r) denote the functional dependence of the force on the distance along the radius vector. If \mathbf{F} is assumed to be directed along \mathbf{r} , we have $\mathbf{F} = F(r)\mathbf{r}/r$, whence the components of the force along the x and y axes become

$$F_{x} = F(r) \cos \theta,$$

$$F_{y} = F(r) \sin \theta,$$
(3.6-1)

where θ is the angle which the radius vector makes with the x axis. But

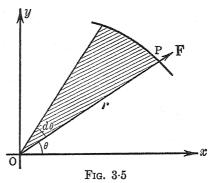
$$\cos \theta = \frac{x}{r}, \quad \sin \theta = \frac{y}{r},$$
 (3.6-2)

whence the equations take the form

$$m\ddot{x} = F(r) \frac{x}{r},$$

$$m\ddot{y} = F(r) \frac{y}{r}.$$
(3.6-3)

It is to be noted that if F(r) is positive the central force is one of repulsion, while if F(r) is negative, the force is attractive. Before we integrate these equations for special cases, we can gain



much information of a general nature about motion in a central force field. Multiply the first equation in (3.6-3) by y and the second by x and subtract. We then have

$$x\ddot{y} - y\ddot{x} = 0. \tag{3.6-4}$$

This can be integrated at once and yields

$$x\dot{y} - y\dot{x} = k, \qquad (3.6-5)$$

where k is an arbitrary constant. The physical significance of eq. (3.6–5) will be better understood if we transform from rectangular to polar coördinates r and θ , as given by the defining equations (3.6–2). Thus

$$\dot{x} = \dot{r}\cos\theta - r\sin\theta \cdot \dot{\theta},
\dot{y} = \dot{r}\sin\theta + r\cos\theta \cdot \dot{\theta},$$
(3.6-6)

whence (3.6-5) becomes at once

$$r^2\dot{\theta} = k. (3.6-7)$$

Referring again to Fig. 3.5, note that the area included between the path of the particle and the radii vectores r and r + dr (enclosing the central angle $d\theta$) is equal to $\frac{1}{2}r^2d\theta$. Therefore eq. (3.6–7) expresses the important fact that a particle in motion under the influence of a central force, irrespective of the exact form of F(r), moves in such a way that the radius vector traces out equal areas in equal times. This is often called the law of areas. We shall have further occasion to note its significance when we come to the study of planetary motion. The constant k (the area constant) is numerically equal to twice the area swept out in unit time. Its value will depend on the initial circumstances of the motion.

To carry the discussion further it will be wise to employ polar coördinates exclusively. From (3.6-6) we have, differentiating again with respect to the time,

$$\ddot{x} = (\ddot{r} - r\dot{\theta}^2)\cos\theta - (2\dot{r}\dot{\theta} + r\ddot{\theta})\sin\theta,$$

$$\ddot{y} = (\ddot{r} - r\dot{\theta}^2)\sin\theta + (2\dot{r}\dot{\theta} + r\ddot{\theta})\cos\theta.$$
(3.6-8)

But from (3.6-7) it follows that

$$2\dot{r}\dot{\theta} + r\ddot{\theta} = 0. \tag{3.6-9}$$

Hence (3.6-8) reduce to

$$\ddot{x} = (\ddot{r} - r\dot{\theta}^2) \cos \theta,$$

$$\ddot{y} = (\ddot{r} - r\dot{\theta}^2) \sin \theta.$$
(3.6-10)

This means that the resultant acceleration in central field motion is directed along the radius vector r (a direct result, of course, of the fact that the force is *central*) and has the magnitude

$$a_r = \ddot{r} - r\dot{\theta}^2. \tag{3.6-11}$$

We recognize the second term $r\dot{\theta}^2$ as the centripetal acceleration [cf. e.g. (1.5-12)]. The term \ddot{r} is the second time rate of change of the length of the radius vector. From (3.6-4) it follows that we can write the equations of motion for a central field in the form of a single equation, viz.,

$$m\ddot{r} - mr\dot{\theta}^2 = F(r), \qquad (3.6-12)$$

where $mr\dot{\theta}^2$ will be known as the centripetal force. Let us try to

eliminate the time from (3.6-12) and obtain a differential equation connecting r and θ . From (3.6-7) we can write

$$r\dot{\theta}^2 = k^2/r^3. {(3.6-13)}$$

It is now simpler to introduce the transformation

$$r = \frac{1}{v} {(3.6-14)}$$

whence

$$\dot{r} = \frac{-\dot{u}}{u^2} = -\frac{du}{d\theta} \cdot \frac{\dot{\theta}}{u^2} = -k \frac{du}{d\theta},$$

and

$$\ddot{r} = -k^2 u^2 \frac{d^2 u}{d\theta^2}.$$

Finally eq. (3.6-12) becomes by substitution of these results

$$\frac{d^2u}{d\theta^2} + u + \frac{f(u)}{mk^2u^2} = 0, (3.6-15)$$

where f(u) = F(r), i.e., when r is set equal to 1/u in F(r) the result is f(u). Eq. (3.6–15) connects u with θ and hence is the differential equation of the orbit of the particle in the central field. Its integration will yield the path directly. It will be noted that our procedure here is different from that pursued in Secs. 3.3 and 3.5 where we integrated the component equations of motion separately and obtained parametric equations for the path. Here we have followed essentially the method of Sec. 3.4 and have eliminated the time from the differential equation of motion before integrating. This course is indeed dictated by the fact that in general the component equations of motion (3.6–3) cannot be integrated separately.

The solution of eq. (3.6-15) will involve two arbitrary constants. A third is already represented by the area constant k. We need one more since in this problem we are really integrating two second order ordinary differential equations. The fourth appears in the integration of (3.6-7), after the result of integrating (3.6-15) (i.e., r in terms of θ) has been substituted into (3.6-7). After this substitution, eq. (3.6-7) becomes the "time" equation of the motion, since it gives dr or $d\theta$ in terms of dt and, on integration, tells how r and θ depend on t. The four arbitrary constants are determined as usual by the initial conditions. We shall illustrate

this by the special cases discussed in the remaining sections of this chapter.

3.7. Motion in a Central Force Field Varying Inversely as the Square of the Distance. We shall now discuss the most important illustration of a central force, namely that in which the force function

$$F(r) = -\frac{c}{r^2}, \qquad (3.7-1)$$

c being a positive constant, corresponding to attraction of the particle to the force center. The case of repulsion will be discussed later. The fundamental equation (3.6-15) now becomes

$$\frac{d^2u}{d\theta^2} + u - \frac{c}{mk^2} = 0, (3.7-2)$$

which is most simply integrated by placing $w = u - c/mk^2$. Then (3·7-2) becomes

$$\frac{d^2w}{d\theta^2} + w = 0. ag{3.7-3}$$

We have already encountered an equation of this mathematical form in Sec. 2.2 (with t as the independent variable). Hence the solution may be written at once

$$u = A\cos(\theta - \alpha) + C, \tag{3.7-4}$$

where $C=c/mk^2$ and A and α are the new arbitrary constants. Reverting to r we have

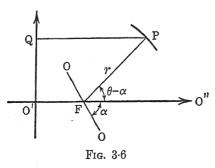
$$r = \frac{1}{A\cos(\theta - \alpha) + C}.$$
 (3.7-5)

This is the polar equation of a conic section, i.e., the locus of a point whose distances from a fixed point and a fixed line are in a constant ratio. Let us write (3.7–5) in the form

$$r = \frac{1/C}{1 + \frac{A}{C}\cos(\theta - \alpha)}.$$
 (3.7-6)

¹ See, for example, Smith and Gale, *loc. cit.*, p. 185, or F. D. Murnaghan, "Analytic Geometry" (Prentice-Hall, N. Y., 1946), p. 307.

Now referring to the figure (Fig. 3.6), let F be the given fixed point on the so-called polar axis OO, and let r be the distance from the particle P to F, with FP making the angle $\theta - \alpha$ with FO'', where the line O'O'' makes the angle α with the polar line. The angle θ is,



of course, the angle the radius vector makes with the polar axis. Let O'Q, perpendicular to the line O'O'', be the given fixed line and O'F = -1/A. Then eq. (3.7-6) expresses the fact that the ratio of r to the distance

$$PQ = [O'F + r\cos(\theta - \alpha)]$$

has the constant value -A/C. This ratio is defined as the eccentricity ϵ of the conic, while the distance O'F is usually written p. The point F is called the focus. In this notation the equation becomes

$$r = \frac{\epsilon p}{1 - \epsilon \cos (\theta - \alpha)} \cdot \tag{3.7-7}$$

It is shown in texts on analytic geometry that for $\epsilon=1$, the conic is a parabola; for $\epsilon<1$, the conic is an ellipse; while for $\epsilon>1$, the conic is a hyperbola. The exact shape of the orbit thus depends on the initial conditions, as we emphasized in the previous section. Since $\epsilon=-A/C$ and $C=c/mk^2$, we have

$$\epsilon = -mk^2 \frac{A}{c}, \qquad (3.7-8)$$

where since c and ϵ are positive, A must be a negative constant. Thus according as

$$-A \leq \frac{c}{mk^2} \tag{3.7-9}$$

there results an elliptic, parabolic or hyperbolic orbit. The physical considerations underlying the distinction between the various types of orbits are more clearly brought out by a study of the initial conditions.

First let us notice from eq. (3.6-6) that the magnitude of the resultant velocity of the particle is

$$v^2 = \dot{x}^2 + \dot{y}^2 = \dot{r}^2 + r^2 \dot{\theta}^2, \tag{3.7-10}$$

where \dot{r} appears as the component of the velocity along the radius vector, and $r\dot{\theta}$ the component of the velocity perpendicular to the radius vector. (Cf. eq. (1·4–3) for the latter for the special case of a circle. It is, however, true for any continuous plane motion.) From eq. (3·7–7) we obtain by time differentiation

$$\dot{r} = -\frac{\epsilon^2 p \sin (\theta - \alpha) \dot{\theta}}{1 - \epsilon \cos^2 (\theta - \alpha)} = -\frac{k \sin (\theta - \alpha)}{p}, \quad (3.7-11)$$

where we have also used (3.7-7) again as well as the law of areas (3.6-7). Finally

$$v^{2} = \frac{k^{2} \sin^{2} (\theta - \alpha)}{p^{2}} + \frac{k^{2}}{r^{2}}.$$
 (3.7-12)

If now for r we substitute once more from (3.7-7) the result is, after a certain amount of algebraic manipulation,

$$\epsilon^2 - 1 = \frac{\epsilon^2 p^2}{k^2} \left(v^2 - \frac{2k^2}{\epsilon pr} \right) \cdot \tag{3.7-13}$$

But in eq. (3.7-7)

$$\epsilon p = -\frac{A}{C} \cdot \left(-\frac{1}{A}\right) = \frac{1}{C} = \frac{mk^2}{c}. \tag{3.7-14}$$

Hence we may rewrite (3.7-13) in the form

$$\epsilon^2 - 1 = \frac{m^2 k^2}{c^2} \left(v^2 - \frac{2c}{mr} \right).$$
 (3.7-15)

Now suppose that the particle is initially projected with velocity v_0 from the initial position $r = r_0$. We still have

$$\epsilon^2 - 1 = \frac{m^2 k^2}{c^2} \left(v_0^2 - \frac{2c}{mr_0} \right).$$
 (3.7–16)

If $v_0 = \sqrt{2c/mr_0}$, $\epsilon = 1$ and the orbit is a parabola. If $v_0 < \sqrt{2c/mr_0}$, $\epsilon < 1$ and the orbit is an ellipse. Finally if $v_0 > \sqrt{2c/mr_0}$, $\epsilon > 1$ and the orbit is a hyperbola.

The velocity $\sqrt{2c/mr_0}$ has an interesting physical significance.

Returning to (3.7-15), we note that if the particle is considered to start at rest at an infinite distance from the force center, i.e., v = 0 for $r = \infty$, the left side of (3.7-15) must vanish identically. Hence $\epsilon = 1$, and the path is a degenerate parabola (the straight line joining the initial position to the force center). But this means that

$$v^2 = \frac{2c}{mr} \tag{3.7-17}$$

gives the velocity attained in the motion at the distance r. Hence at distance r_0 the velocity attained is

$$v = \sqrt{\frac{2c}{mr_0}} \cdot \tag{3.7-18}$$

The results at the end of the previous paragraph may then be restated: the orbit in an attractive inverse square force field is a parabola, ellipse, or hyperbola according as the velocity of projection at r_0 is equal to, less than, or greater than the velocity attained at r_0 by a particle falling freely in this field from infinity.

It is worthy of comment that, although the kind of conic represented by the orbit depends only on the speed v_0 , the actual eccentricity ϵ for elliptic and hyperbolic orbits still depends on the area constant k and is therefore dependent on the direction of projection as well as the speed.

At the end of Sec. 3.6 we mentioned the "time" equation which expresses r or θ in terms of the time t. We obtain it in general by the use of eq. (3.6–7) and the equation of the orbit, (3.7–7). Thus from the latter we have

$$\theta - \alpha = \arccos\left(\frac{1}{\epsilon} - \frac{p}{r}\right),$$
 (3.7-19)

so that differentiating with respect to the time

$$\dot{\theta} = \frac{\pm p\dot{r}}{r^2 \sqrt{1 - \left(\frac{1}{\epsilon} - \frac{p}{r}\right)^2}} = \frac{k}{r^2} \cdot \tag{3.7-20}$$

Hence we have for the determination of r in terms of t, the differential equation

$$\dot{r} = \pm \frac{k}{p} \sqrt{1 - \left(\frac{1}{\epsilon} - \frac{p}{r}\right)^2} \cdot \tag{3.7-21}$$

The integration of eq. (3.7-21) for the special case of planetary elliptic motion will be given in the next section.

3.8. Planetary Motion. About the beginning of the seventeenth century the laws of planetary motion were enunciated by Kepler after many years of laborious calculation and reduction of observations of the positions of the planets. These laws are three in number, viz. (1) that the planets describe, relatively to the sun, ellipses of which the sun occupies a focus; (2) the radius vector of each planet traces out equal areas in equal times; (3) the squares of the periodic times (i.e., periods of revolution) of any two planets vary as the cubes of the major axes of their orbits. These laws do not imply that the sun must be stationary in the reference system being used. In this section we shall actually make this assumption, reserving the general case for later consideration.

We recall at once that the second law is satisfied by any particle moving in a central field of force [eq. (3.6-7)]. It remains to be shown, however, that if the second law is satisfied the force acting on the planet is directed toward the sun, i.e., it is a central force. But this follows at once from the result of differentiating eq. (3.6-5) (the expression of the second law in rectangular coordinates) with respect to the time. The result is, of course

$$\frac{\ddot{y}}{\ddot{x}} = \frac{y}{x},\tag{3.8-1}$$

whence, since $m\ddot{y}$ and $m\ddot{x}$ are respectively the y and x components of the force, the latter must be directed through the origin and hence be a central force.

We can next show that the first law of Kepler implies an inverse square law of force. The polar equation of an ellipse referred to the focus as pole is (with $\alpha = 0$ for convenience)

$$\frac{1}{r} = u = \frac{1 - \epsilon \cos \theta}{\epsilon p}, \qquad (3.8-2)$$

as given in eq. (3.7-7). We therefore have

$$\frac{d^2u}{d\theta^2} = \frac{\cos\theta}{p} \cdot \tag{3.8-3}$$

Now for any central force we have shown that the function f(u) [eq. (3.6-15)] obeys the equation

$$f(u) = -mk^2u^2\left(u + \frac{d^2u}{d\theta^2}\right).$$

On substitution from (3.8-2) and (3.8-3) we therefore have

$$f(u) = -\frac{mk^2u^2}{\epsilon p}, \qquad (3.8-4)$$

which shows that the force varies inversely as the square of the distance, the negative sign of course indicating attraction.

The third law of Kepler may also be shown to follow from the laws of motion in an inverse square central force field. For we have

$$r^2\dot{\theta} = 2\dot{S} = k, \tag{3.8-5}$$

where S is the area swept out in time t by the radius vector. In one period P of the motion the area traversed will be the area of the ellipse itself, or πab , where a is the semi-major axis and b is the semi-minor axis. (That this is the area may be verified by simple integration.) Noting the relation¹

$$b = a\sqrt{1 - \epsilon^2},\tag{3.8-6}$$

we have

$$P = \frac{2\pi a^2}{k} \sqrt{1 - \epsilon^2}.$$
 (3.8–7)

From eq. (3.8-2) in connection with Fig. 3.7 we see that

$$r_{\text{max}} = \frac{mk^2}{c(1-\epsilon)}$$
, and $r_{\text{min}} = \frac{mk^2}{c(1+\epsilon)}$, (3.8-8)

corresponding respectively to the positions marked A and B. The former is termed aphelion while the latter is the perihelion. Clearly

$$2a = r_{\text{max}} + r_{\text{min}}$$

$$= \frac{mk^2}{c} \cdot \frac{2}{1 - \epsilon^2} \cdot \tag{3.8-9}$$

¹ Smith and Gale, loc. cit., p. 162.

Therefore

$$\sqrt{1-\epsilon^2}=\sqrt{\frac{mk^2}{ca}},$$

and (3.8-7) becomes

$$P^2 = \frac{4\pi^2 ma^3}{c}, (3.8-10)$$

which is the mathematical statement of Kepler's third law.

It will be recalled that the connection between the laws of

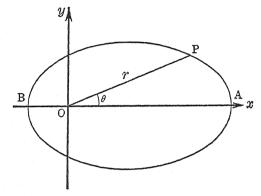


Fig. 3.7

Kepler and motion in an inverse square central force field led Newton ultimately to propound his law of universal gravitation already stated and discussed in Sec. 2·3.

The quantity c, the coefficient in the central force expression used in this and the previous section, can be expressed in terms of the gravitational constant G (see Sec. 2.3) if the central force is due to a particle of mass M. Then c = GmM, and the equation (3.8–10) can be written in the form

$$P^2 = \frac{4\pi^2 a^3}{GM} \cdot (3.8-11)$$

From this it follows that if a planet has a satellite whose period of revolution P about the planet can be observed as well as the semi-major axis a of its orbit, the mass of the planet may thereby be obtained. The mass of the sun is, of course, obtainable by the use of the same equation.

It is interesting to note that the orbits of many comets (the non-returning kind) have been found to be parabolas, a special variety of central field motion, thus lending further support to Newton's law.

The "time" equation for an elliptical planetary orbit can be obtained by the integration of eq. (3.7–21) of the previous section. The latter becomes on separation of variables and choosing the plus sign

$$\frac{dr}{\sqrt{1 - \left(\frac{1}{\epsilon} - \frac{p}{r}\right)^2}} = \frac{k}{p} dt. \tag{3.8-12}$$

If we recall that the quantity p is given by

$$p = \frac{mk^2}{c\epsilon}, \qquad (3.8-13)$$

we have from (3.8-9) the relation

$$\epsilon p = a(1 - \epsilon^2). \tag{3.8-14}$$

Substituting for p in terms of a and ϵ in eq. (3.8–12), we are led to the form

$$\frac{r\,dr}{\sqrt{\epsilon^2 a^2 - (a-r)^2}} = \frac{k}{a\sqrt{1-\epsilon^2}}\,dt. \tag{3.8-15}$$

Now it is of advantage to introduce the observable quantity P, the period of the planet's revolution. From eq. (3.8–10) we have

$$\frac{1}{P} = \frac{1}{2\pi} \sqrt{\frac{c}{ma^3}}.$$

But

$$\frac{c}{m} = \frac{k^2}{\epsilon p} = \frac{k^2}{a(1 - \epsilon^2)},\tag{3.8-16}$$

from eq. (3.8-13) above. Therefore

$$\frac{1}{P} = \frac{1}{2\pi} \cdot \frac{k}{a^2 \sqrt{1 - \epsilon^2}},\tag{3.8-17}$$

and eq. (3.8-15) becomes

$$\frac{r}{a} \cdot \frac{dr}{\sqrt{\epsilon^2 a^2 - (a-r)^2}} = \frac{2\pi}{P} dt.$$
 (3.8–18)

To integrate this equation conveniently it is customary to introduce a new variable E defined by the equation

$$a - r = a\epsilon \cos E, \tag{3.8-19}$$

so that

to

$$r = a(1 - \epsilon \cos E). \tag{3.8-20}$$

E is an angle and is called in texts on celestial mechanics the eccentric anomaly. To understand the geometrical significance of the latter, let us refer to the following figure (Fig. 3.8). This

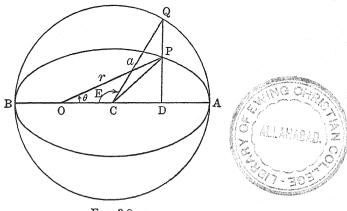


Fig. 3.8

presents the elliptical orbit APB with the focus at O and center at C. Let us draw the circle AQB of radius a about C as center. This is termed the *auxiliary* circle. Suppose the planet is at the position P with coördinates r and θ . Through P construct PD, the perpendicular to the major axis AB, and let it intersect the circle at Q. Then the angle QCB = E, the eccentric anomaly. That this satisfies the definition given in the relation (3-8-20) may be seen from the fact that from the figure we have

$$r\cos\theta + a\cos E = OC. \tag{3.8-21}$$

But $OC = a - r_{\min} = a - \frac{mk^2}{c(1+\epsilon)}$ from (3.8–8), and this reduces

$$OC = \frac{a + a\epsilon - \epsilon p}{1 + \epsilon} = \frac{a + a\epsilon - a + a\epsilon^2}{1 + \epsilon}$$

Hence (3.8-21) becomes, using eq. (3.8-2),

$$\frac{r}{\epsilon} - p + a \cos E = a\epsilon,$$

or

$$r = \epsilon p + a\epsilon^2 - a\epsilon \cos E. \tag{3.8-22}$$

But $\epsilon p = a(1 - \epsilon^2)$ and hence (3.8-20) follows.

Now substituting into (3.8-18) the differential equation becomes

$$\frac{2\pi}{P} dt = (1 - \epsilon \cos E) dE, \qquad (3.8-23)$$

and integrating we have

$$\frac{2\pi}{P}(t-t_0) = E - \epsilon \sin E, \qquad (3.8-24)$$

where t_0 is the time for E=0, i.e., the passage through perihelion (see the figure). Now $2\pi/P \cdot (t-t_0)$ represents the angle through which the radius vector would have moved at the uniform angular velocity $\omega=2\pi/P$. This angle is called in celestial mechanics the *mean anomaly* and often denoted by M. The relation

$$M = E - \epsilon \sin E \tag{3.8-25}$$

is known as *Kepler's equation*, and as soon as M is known for any time it can be solved for E, whence r and θ can be found for any instant. However, since the equation (3.8–25) is transcendental, its solution cannot be expressed in a closed form. Methods for its approximate solution in various cases will be found in texts on celestial mechanics.¹

3.9. Electron Motions in the Bohr Model of the Hydrogen Atom. The Bohr-Rutherford atom model, which is still of great value in pictorial representations of atomic structure, pictures the hydrogen atom as formed of a heavy nucleus with a charge of positive electricity of magnitude $e=4.80\times10^{-10}$ electrostatic units, and a negatively charged electron with charge -e. To a first approximation the electron is supposed to move about the nucleus

¹ For example, see F. R. Moulton, "Celestial Mechanics" (Macmillan, New York, 2nd ed., 1914), p. 160 ff.

in an orbit due to the central electrostatic attraction of magnitude $-e^2/r^2$, where r is the distance between the particles. The gravitational attraction can easily be shown to be negligible compared with the electrostatic, for the mass of the hydrogen nucleus is 1.66×10^{-24} grams while that of the electron is approximately 0.9×10^{-27} grams. As a matter of fact the investigation in a later chapter will show that the two particles both move about their common center of mass (see Sec. 6.3). However, the motion of the nucleus due to its much larger mass is so slight compared with that of the electron that for the present we may neglect it and consider that the electron moves in a central field about an approximately stationary nucleus. Since the orbit of the electron must be bounded in space as long as it remains in the atom, the path will then be an ellipse with the nucleus at one focus, as we learn from our study of the analogous case of planetary motion discussed in Sec. 3.8.

Now an outstanding feature of the Bohr theory is the postulate that not all elliptical orbits are possible for the electron, i.e., not all values of the semi-major axis a and eccentricity ϵ occur. The possible orbits are those for which the following extra-mechanical conditions are satisfied, viz.,

$$\oint p_{\theta} d\theta = n_1 h,$$

$$\oint p_r dr = n_2 h.$$
(3.9-1)

These relations are to be interpreted thus: p_{θ} is the so-called moment of momentum of the electron in its path, viz.,

$$p_{\theta} = mr^2 \dot{\theta} = mrv_1, \tag{3.9-2}$$

where $v_1 = r\dot{\theta} = \text{velocity component at any instant perpendicular}$ to the radius vector [recall eq. (1.4-3)]. On the other hand p_r is the so-called radial momentum, i.e.,

$$p_r = m\hat{r}. \tag{3.9-3}$$

¹ The moment of momentum is now commonly referred to as the *angular momentum*, though the latter term is more strictly applicable to rigid bodies than to single particles. Note that p_{θ} is also interpretable as the product of mass and area constant [cf. eq. (3·6–7)].

The form in which the integrals in (3.9-1) are written is to indicate that the integration is to extend over the whole region of variation of the variables in question. Thus the first integration will extend from $\theta = 0$ to $\theta = 2\pi$ as limits, while the second is taken from r_{\min} (i.e., perihelion) to r_{\max} (aphelion) and back again to r_{\min} along the elliptical path. The actual method of integration will be indicated below. The quantity h is the fundamental constant of the quantum theory, the so-called constant of Planck.1 The quantities n_1 and n_2 are integers, and the conditions expressed in the equations (3.9-1) are known as the quantum conditions. We must emphasize that they are not deductions from previous theory but are outright postulates or assumptions. We cannot here be concerned with their true significance for it would take a considerable discussion of atomic structure theory to make this clear. What we are interested in is their influence on the possible motions of the electron.

Substituting the value of p_{θ} from (3·9–2) into the first condition and recalling that for a central field motion $r^2\dot{\theta}=k$ [eq. (3·6–7)] we have:

$$\int_0^{2\pi} mk \ d\theta = n_1 h,$$

or

$$mk = p_{\theta} = \text{const.} = \frac{n_1 h}{2\pi} \cdot$$
 (3.9-4)

The carrying out of the integration in the second condition is somewhat more troublesome. Let us write the polar equation of the ellipse in the form

$$r = \frac{a(1 - \epsilon^2)}{1 - \epsilon \cos \theta},\tag{3.9-5}$$

¹ The student's attention at this point is directed to the good popular treatment of the classical quantum theory of atomic structure in A. Haas, "The World of Atoms" (D. Van Nostrand, New York, 1929). See also in the same connection the readable work of K. K. Darrow, "Introduction to Contemporary Physics" (D. Van Nostrand, New York, 1926) or the more recent text by G. E. M. Jauncey, "Modern Physics" (D. Van Nostrand, New York, 1932). The more ambitious student may profitably consult Ruark and Urey, "Atoms, Molecules, and Quanta" (McGraw-Hill, New York, 1930), Loeb, "Atomic Structure" (Wiley, New York, 1938), or H. Semat, "Introduction to Atomic Physics" (Rinehart, New York, 2d ed., 1946).

which follows from eq. (3.8-2) if we note that

$$\epsilon p = \frac{mk^2}{c} = a(1 - \epsilon^2), \qquad (3.9-6)$$

from eqs. (3.7-14) and (3.8-9). Now the radial momentum can be transformed as follows

$$p_r = mr = m\frac{dr}{d\theta}\dot{\theta} = \frac{mk}{r^2}\frac{dr}{d\theta}, \qquad (3.9-7)$$

while

$$dr = \frac{dr}{d\theta} d\theta = -\frac{a\epsilon(1 - \epsilon^2)\sin\theta d\theta}{(1 - \epsilon\cos\theta)^2},$$
 (3.9-8)

from eq. (3.9-5). Therefore on substitution the second condition becomes

$$mk\epsilon^2 \int_0^{2\pi} \frac{\sin^2 \theta}{(1 - \epsilon \cos \theta)^2} d\theta = n_2 h.$$
 (3.9-9)

We integrate this by parts, obtaining

$$n_{2}h = mk\epsilon \left\{ -\left[\frac{\sin\theta}{1 - \epsilon\cos\theta}\right]_{0}^{2\pi} + \int_{0}^{2\pi} \frac{\cos\theta}{1 - \epsilon\cos\theta} d\theta \right\},\,$$

which reduces to

$$n_2 h = mk \int_0^{2\pi} \left(\frac{1}{1 - \epsilon \cos \theta} - 1 \right) d\theta, \qquad (3.9-10)$$

since the square bracket above vanishes on the introduction of the limits. The integral

$$\int_0^{2\pi} \frac{d\theta}{1 - \epsilon \cos \theta} = \frac{2\pi}{\sqrt{1 - \epsilon^2}},$$
 (3.9-11)

from No. 300 in Peirce's Table of Integrals. Therefore the second quantum condition becomes

$$\frac{1}{\sqrt{1-\epsilon^2}} - 1 = \frac{n_2 h}{2\pi k m}$$

$$= \frac{n_2}{n_1}, \qquad (3.9-12)$$

utilizing the first condition [eq. (3.9-4)], whence

$$\epsilon^2 = 1 - \frac{n_1^2}{(n_1 + n_2)^2}$$
 (3.9–13)

But this means that the possible orbits of the electron are restricted to those for which the eccentricity ϵ takes the values allowed by (3.9–13), as n_1 and n_2 assume all possible positive in-

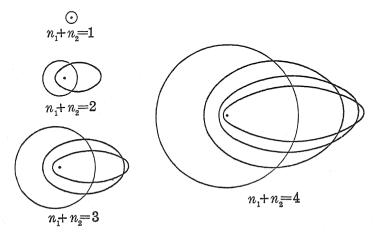


Fig. 3.9

tegral values.¹ To each value of ϵ correspond also definite major and minor axes. Thus combining (3.9-4) and (3.9-6) we have

$$a = \frac{(n_1 + n_2)^2 h^2}{4\pi^2 m e^2}, \qquad (3.9-14)$$

noting that $c = e^2$ in this special problem. The dimensions of the allowed orbits are thus fully fixed by the quantum conditions. A few sketches for certain selected small values of n_1 and n_2 will serve to illustrate the essential results we have obtained. (Fig. 3.9.)

The times of revolution in the allowed orbits are given at once from Kepler's third law, eq. (3.8-10). Thus substituting into

¹ It is customary to exclude the value $n_1=0$ since this corresponds to $\epsilon=1$ and a straight line orbit through the nucleus, which is physically meaningless.

(3.8-10) from (3.9-14) we have

$$P = \frac{(n_1 + n_2)^3 h^3}{4\pi^2 m e^4},$$
 (3.9–15)

the periodic times thus increasing with the cubes of the natural numbers.

It may also be remarked that under certain conditions an electron coming from outside the atom will move in a hyperbolic path about the nucleus. These conditions are, indeed, precisely those laid down in Sec. 3.7.

Further discussion of the application of the Bohr theory will be found in Sec. 4.4.

3.10. Motion in a Repulsive Inverse Square Force Field. Alpha Particle Deflection. In our discussion of the motion of a particle in an inverse square field we have dealt wholly with the case where the force is attractive. There is, however, an important significance to be attached to the case in which the force is repulsive. Consider a positively charged nucleus of an atom (Fig. 3.10)

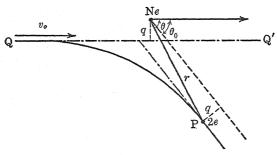


Fig. 3.10

with charge Ne (N being an integer and e the fundamental unit of charge, 4.80×10^{-10} e.s. units). Next suppose that an alpha particle, a doubly charged helium atom, i.e., the nucleus of a helium atom with charge 2e, is projected into the neighborhood of the nucleus Ne in such a way that at a very great distance where the repulsive action is very slight the alpha particle has a velocity v_0 in the direction QQ'. For convenience we shall describe the motion of the alpha particle in terms of polar coördinates r, θ , choosing the polar axis parallel to QQ' and placing Ne at the pole. The

perpendicular distance between QQ' and the axis is denoted by q. The force of repulsion between the alpha particle and the nucleus is of magnitude $2Ne^2/r^2$, and therefore the analysis of Sec. 3.7 will at once apply, with the quantity c of that section now assuming the value $-2Ne^2$, i.e., being negative instead of positive. The orbit of the alpha particle will therefore be a conic with equation [see eq. (3.7-7)]

$$r = \frac{\epsilon p}{1 - \epsilon \cos(\theta - \alpha)} \cdot \tag{3.10-1}$$

From eq. (3·7-16) we see that since $v_0 > 0$ for $r_0 = \infty$, $\epsilon > 1$ and the orbit is a hyperbola. As Fig. 3·10 shows, r approaches ∞ for two values of θ , namely π and θ_0 . Hence we have from (3·10-1) the conditions

$$\cos (\pi - \alpha) = \frac{1}{\epsilon} = -\cos \alpha, \qquad (3.10-2)$$

and

$$\cos (\theta_0 - \alpha) = \frac{1}{\epsilon} \cdot \tag{3.10-3}$$

Combining (3·10-2) and (3·10-3) yields (since $\sin \alpha = \sqrt{\epsilon^2 - 1/\epsilon}$)

$$\frac{1+\cos\theta_0}{\sin\theta_0}=\sqrt{\epsilon^2-1},\qquad (3\cdot10-4)$$

or in terms of $\theta_0/2$,

$$\cot\frac{\theta_0}{2} = \sqrt{\epsilon^2 - 1}. (3.10-5)$$

Now if we examine Fig. 3.10 we see that as $\theta \to \pi$, we have to an increasingly good approximation

$$r\sin\theta \doteq q. \tag{3.10-6}$$

Hence in the neighborhood of $\theta = \pi$

$$\dot{r} \doteq -q \cos \theta \cdot \frac{\dot{\theta}}{\sin^2 \theta} \doteq -r\dot{\theta} \sqrt{\frac{r^2}{q^2} - 1}. \qquad (3.10-7)$$

Applying (3.7-10) we are led at once to the simple but important result

$$k^2 = v_0^2 q^2. (3.10-8)$$

Now return to (3.7-13) which becomes for $v = v_0$ at $r = \infty$

$$\epsilon^2 - 1 = \frac{\epsilon^2 p^2 v_0^2}{k^2}$$
 (3·10–9)

Employing (3.7–14) to substitute for ϵp , we get

$$\epsilon^2 - 1 = \frac{m^2 k^2 v_0^2}{c^2} = \frac{m^2 k^2 v_0^2}{4N^2 e^4} \cdot$$
 (3·10–10)

Using (3.10-8) we obtain finally

$$\sqrt{\epsilon^2 - 1} = \frac{mv_0^2 q}{2Ne^2}, \tag{3.10-11}$$

whence (3.10-5) becomes

$$\cot \frac{\theta_0}{2} = \frac{m v_0^2 q}{2Ne^2} \cdot \tag{3.10-12}$$

This will give us the angular deviation of the alpha particle from its original path. Let us ascertain its order of magnitude in a special case. For an alpha particle we have approximately $m=6.6\times 10^{-24}$ grams, and v_0 may be taken as approximately 2×10^9 cm/sec, which is very close to the actual value for the velocity of α -particles from Radium C. The velocities of these particles depend on the source. The electronic charge e has already been given as 4.80×10^{-10} electrostatic units. Substitution then yields

$$\cot \frac{\theta_0}{2} = 5.83 \times 10^{13} \frac{q}{N}$$
 (3.10–13)

Let us suppose that N=79, i.e., that the nucleus is that of gold. In order to produce a deflection of 150°, such as has been experimentally observed in the case of gold, the α -particle must approach the nucleus within such a distance that $q=3.66\times 10^{-13}$ cm approximately. It is evident that experiments on the deflection of alpha particles are of great value in enabling us to estimate

¹ These data are taken from Birge, "Values of the General Physical Constants," *Reviews of Modern Physics*, July 1929. For the methods of determining them and for other information on atomic theory the student is recommended to read the books referred to in Sec. 3.9.

² See Lord Rutherford, Chadwick and Ellis, "Radiations from Radioactive Substances" (Cambridge University Press, 1930), p. 47.

limits for the dimensions of the nucleus. As a matter of fact it should be noted that the whole nuclear theory of atomic structure rests on experiments of this nature (with the accompanying theory) carried out by Rutherford about 1911.

PROBLEMS

- 1. A projectile is shot off with an initial velocity of 2500 ft/sec at an angle of 30° to the horizontal. Assuming no air resistance calculate the total time spent in the air, the range, and the total length of the path in the air.
- 2. A projectile travels in a medium which resists its motion with a force of the form $k\mathbf{v}$, where k is a coefficient of proportionality and \mathbf{v} is the vector velocity in the orbit, with components \dot{x} and \dot{y} respectively. Calculate the range and time spent in air as a function of k and the projection velocity and angle.
- 3. A projectile is projected at angle α with velocity v_0 . Find the time at which the angle which the direction of motion makes with the direction of projection is $\pi/2 \alpha$.
- 4. Two simple harmonic motions take place along the x and y axes respectively with equal amplitudes but with frequencies in the ratio 4:3. Find the curve resulting from the composition of the two motions.
- 5. Calculate approximately the mass of the earth and the mass of the sun from Newton's law of gravitation. Data: radius of earth, 4000 miles; g=980 cm/sec²; G, the constant of gravitation, $=6.67\times 10^{-8}\,\frac{\rm dyne~cm^2}{\rm gm^2}$; $R={\rm mean}$ distance from earth to sun = 93,000,000 miles.
- 6. Calculate the orbit of a particle moving in a plane under the influence of a central force varying directly as the first power of the distance from the force center.
- 7. Find the way in which the central force must vary with the distance from the force center in order that the particle may describe the spiral $r=1/c\theta$, where c is constant and r and θ are the usual polar coördinates.
- 8. Find the way in which the central force must vary with the distance from the force center in order that a particle may describe the lemniscate

$$r^2 = a^2 \cos 2\theta,$$

the force center being assumed to be at the pole. Use the law of areas to find the expression for the resultant velocity of the particle as a function of r.

- 9. The expression $2r\dot{\theta}$ in central field motion is usually called the Coriolis acceleration. Show that it is equal to $-r\dot{\theta}$. Hence find the expression for $\dot{\theta}$ in an elliptical inverse square orbit as a function of θ .
 - ¹ See Rutherford, Chadwick and Ellis, loc. cit., Chap. VIII.

- 10. The fourth satellite of Jupiter has a period of approximately 16.75 days and its semi-major axis is approximately 1.88×10^9 meters. Find an approximate value for the mass of the planet.
- 11. The eccentricity of the orbit of the planet Mercury is approximately 0.206. Its period of revolution is 0.24 year. If its semi-major axis is approximately 58×10^6 kilometers, at what rate does its radius vector sweep out area?
- 12. An electron moves in the xy plane under the influence of a uniform electric field directed along the y axis and a uniform magnetic field directed along the z axis (perpendicular to the xy plane). From the theory of electricity and magnetism the component equations of motion are:

$$m\ddot{x} = -e\dot{y}H/c$$

$$m\ddot{y} = -eE + e\dot{x}H/c,$$

where e is the charge on the electron, and E and H are the electric and magnetic field intensities respectively, and c is the conversion factor 3×10^{10} . Integrate to show that the orbit of the electron is a cycloid. Show that one gets different forms of cycloids depending on the choice of initial conditions.

- 13. Plot to scale the various types of orbits allowed to the electron in the hydrogen atom (discussed in the text) by the quantum conditions for the cases where the sum of the two quantum numbers $n_1 + n_2$ is equal to 1, 2, 3, 4 successively. Calculate in each case the perihelion and aphelion distances (i.e., minimum and maximum distances from the nucleus).
- 14. Prove that when a particle moves in any central field of force the velocity at any point is inversely proportional to the perpendicular from the force center to the tangent to the path at that point.
- 15. Under what central force will a particle describe a circle with the force center on the circumference of the circle?

CHAPTER IV

ENERGY IN PARTICLE DYNAMICS

4.1. Potential Energy and the Energy Concept. Conservative Forces. In Chapter I we introduced the concepts of work and kinetic energy (Sec. 1·10). In the discussion of Chapters II and III we made relatively little use of these concepts, restricting ourselves almost completely to force and the equation of motion. It is important to recognize that this is sufficient for a large part of mechanics. However, it is now time that we began to avail ourselves of the powerful concept of energy. We approach this concept in the following way. In Sec. 1·10 we have seen that when a particle of mass m is in motion with velocity having magnitude v (with reference to some inertial system), it is convenient to associate with its motion the quantity $\frac{1}{2}mv^2$, the change in which over a given interval represents the work done by the resultant force on the particle during this interval. We have called this quantity the kinetic energy of the particle.

Let us now examine the work-kinetic energy theorem (eq. 1·10-20) more closely. We return to the general equation

$$\int_{\mathbf{r}_0}^{\mathbf{r}_1} \mathbf{F} \cdot d\mathbf{r} = \frac{1}{2} m v_1^2 - \frac{1}{2} m v_0^2, \qquad (4.1-1)$$

but for simplicity confine our discussion at first to rectilinear motion along the x axis. The equation in question then becomes

$$\int_{x_0}^{x_1} F_x \, dx = \frac{1}{2} m \dot{x}_1^2 - \frac{1}{2} m \dot{x}_0^2. \tag{4.1-2}$$

Let us now suppose that F_x is an integrable function of x, i.e., there exists a function V(x) such that

$$F_x dx = -dV. (4.1-3)$$

On integration, we have, if we denote $\frac{1}{2}m\dot{x}_1^2$ and $\frac{1}{2}m\dot{x}_0^2$ by K_1 and K_0 , respectively,

$$V_1 - V_0 + K_1 - K_0 = 0$$

or

$$K_1 + V_1 = K_0 + V_0 = \text{const.} = U.$$
 (4·1-4)

In other words the quantity U, which is the sum of K and V, remains constant during the motion. We shall call V, which is a function of x, the potential energy of the particle. We have then just shown that for the above simple case if such a function exists the sum of the kinetic and potential energies remains constant. We shall denote this sum (U) as the total mechanical energy of the particle. This is at once subject to generalization to the case of motion in three dimensions. Thus let us write the force vector \mathbf{F} in terms of its components along the x, y and z axes respectively, viz.

$$\mathbf{F} = \mathbf{i}F_x + \mathbf{j}F_y + \mathbf{k}F_z,$$

where F_z , F_y and F_z are, in general, all functions of the three coördinates x, y, z. The displacement vector $d\mathbf{r}$ may also be written likewise

$$d\mathbf{r} = \mathbf{i} \, dx + \mathbf{j} \, dy + \mathbf{k} \, dz.$$

We then have from the definition (1.10-5)

$$\mathbf{F} \cdot d\mathbf{r} = F_x dx + F_y dy + F_z dz,$$

so that eq $(4\cdot1-1)$ becomes, if the integration is conducted from the point P_0 to the point P_1 ,

$$\int_{P_0}^{P_1} (F_x dx + F_y dy + F_z dz) = \frac{1}{2} m v_1^2 - \frac{1}{2} m v_0^2. \quad (4.1-5)$$

Now if the expression $F_x dx + F_y dy + F_z dz$ is a perfect differential, i.e., if there exists a function V(x, y, z) such that

$$F_z dx + F_y dy + F_z dz = -\frac{\partial V}{\partial x} dx - \frac{\partial V}{\partial y} dy - \frac{\partial V}{\partial z} dz = -dV, \quad (4.1-6)$$

we again have

$$K + V = U = \text{const.} \tag{4.1-7}$$

We recall from calculus that dV is called the total differential of the function V(x, y, z). It is the total change in V corresponding to the alterations dx, dy, dz in x, y, z respectively. F_z then appears as the negative rate of change of V with respect to x, viz., $F_x = -\partial V/\partial x$; similarly for F_y and F_z . We are now under the necessity of introducing partial derivative signs since V is a function of more than one variable.

It must be emphasized that the existence of the potential energy V to satisfy (4·1–6) demands a rather stringent condition. It will be recalled from calculus or can be seen without trouble from

$$F_x = -\frac{\partial V}{\partial x}$$
, $F_y = -\frac{\partial V}{\partial y}$, $F_z = -\frac{\partial V}{\partial z}$

that we must have

$$\frac{\partial F_x}{\partial y} = \frac{\partial F_y}{\partial x}, \quad \frac{\partial F_x}{\partial z} = \frac{\partial F_z}{\partial x}, \quad \frac{\partial F_y}{\partial z} = \frac{\partial F_z}{\partial y}. \tag{4.1-8}$$

It will not need much experimentation to convince the reader that very few forces chosen at random will have components satisfying (4·1-8). Thus

$$F_x = ky$$
, $F_y = kz$, $F_z = kx$,

where k is a constant, certainly will not yield a function V. On the other hand

$$F_x = kx$$
, $F_y = ky$, $F_z = kz$

does satisfy (4:1-8) and indeed leads at once to

$$V = \frac{1}{2}k(x^2 + y^2 + z^2).$$

Forces whose components satisfy $(4\cdot1-8)$ and for which the potential energy exists are called *conservative* forces, since for them it is possible to define a total mechanical energy U which remains constant during the motion or may be said to be *conserved*. Eq. $(4\cdot1-7)$ then becomes the celebrated principle of the *conservation of mechanical energy*. The particle in this case is said to constitute a *conservative* system. Since for a conservative system

$$\int_{\mathbf{r}_0}^{\mathbf{r}_1} \mathbf{F} \cdot d\mathbf{r} = -(V_1 - V_0), \qquad (4.1-9)$$

where V_0 is the value of the potential energy at the beginning of the path, and V_1 the value at the end of the path, it follows that the *gain* in potential energy is the negative of the work done by the force during the motion. Or we can say that the gain in potential energy in any motion of a particle under a conservative force must be balanced by an equal loss in kinetic energy.

A simple illustration will help to render our discussion clearer. Revert to the rectilinear motion of a particle attached to a spring (Sec. 2.2). Here the elastic restoring force is

$$F_x = -kx, (4.1-10)$$

where k is the elastic constant.

Hence

$$\int_{x_0}^x F_x dx = -\left(\frac{1}{2}kx^2 - \frac{1}{2}kx_0^2\right). \tag{4.1-11}$$

Therefore we may write for the potential energy of the particle attached to the spring when its displacement is x,

$$V = \frac{1}{2}kx^2 + C, (4.1-12)$$

where C is an arbitrary constant. We see that V is not absolutely determined in that we may always add to it a quantity independent of x. It is desirable to consider V = 0 when x = 0, and therefore in the present case we shall set C = 0. The field of force in the present case is conservative and the total energy of the system is

$$U = \frac{1}{2}mv^2 + \frac{1}{2}kx^2. \tag{4.1-13}$$

It is interesting to observe that this is precisely the form of equation that results from the first integration of the equation of motion of the system

$$m\ddot{x} = -kx, \qquad (4\cdot 1-14)$$

as we recall at once from Sec. 2.2. We can generalize this statement and say that the first integral of the equations of motion of a conservative system is the equation of energy. For let us write the equations of motion in component form

$$m\ddot{x} = F_x, \quad m\ddot{y} = F_y,$$

 $m\ddot{z} = F_z.$ (4·1-15)

Multiplying these respectively by $\dot{x} dt$, $\dot{y} dt$, $\dot{z} dt$, and adding, we obtain

$$d[\frac{1}{2}m(\dot{x}^2+\dot{y}^2+\dot{z}^2)] = F_x dx + F_y dy + F_z dz, \quad (4.1-16)$$

and if the right-hand side is -dV, we have at once on integration

$$\frac{1}{2}mv^2 + V = U = \text{const.}$$
 (4·1-17)

It is thus evident that whenever we are dealing with a conservative system the energy equation will be of considerable importance. Incidentally, examination of $(4\cdot1-17)$ reveals that at the

place where V is a minimum, K (the kinetic energy) is a maximum and vice versa.

4.2. Energy Relations in a Uniform Field. Here the force equation for rectilinear motion is

$$F_x = m\ddot{x} = \text{const.} = k, \tag{4.2-1}$$

and the energy equation is therefore

$$\frac{1}{2}m\dot{x}^2 - kx = U, (4.2-2)$$

the potential energy being V=-kx. This is then the form assumed for the case of a particle falling freely near the surface of the earth. The constant k is in this case -mg, where m is the mass of the particle and g the acceleration of gravity. Suppose a particle dropped from a height h above the ground. Its potential energy at the instant of dropping has the magnitude mgh relative to the earth's surface and at the instant it reaches the earth becomes zero, having all "changed" into kinetic energy. Hence the total energy of the particle will be equal to its potential energy at the start, viz., mgh. We thus have

$$K + V = mgh. (4.2-3)$$

From this the velocity with which it strikes the ground is given by

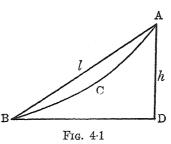
$$v = \sqrt{2gh}, \tag{4.2-4}$$

as in the treatment of Sec. 2.2. The question may arise — what happens to the kinetic energy which the particle has at the moment of striking? The particle may be said to transfer this to the earth as a result of collision. In such a collision heat is developed and sound waves are produced. If we are willing to suppose that these physical manifestations involve energy, we can then say that some at least of the kinetic energy of the particle is "transformed" into energy of another form. We here glimpse the doctrine of the universal conservation of energy, which supposes that mechanical energy as we have defined it cannot be destroyed, but can only be changed into some other kind.

Incidentally we now see another significance of the statement made in Sec. 2.2 concerning the motion of a particle along an

inclined plane. Consulting Fig. 4.1 we recall that the velocity along the plane of length l at the bottom is equal to the velocity which would have been gained by direct vertical fall from A

through the height h. But the meaning of this in terms of the energy equation is now clear. The potential energy is a function of the vertical height x above the level BD of the bottom of the plane. Hence, as long as the motion stays conservative, it makes no difference how the fall takes place to the bottom: the change



in potential energy and the equal change in kinetic energy will be the same. As a matter of fact, this then will still be true if the motion were to take place along the curve ACB. This of course supposes that there is nothing peculiar about this motion such as friction, for example, to render it non-conservative.

4.3. Energy Relations in a Central Force Field. The equations of motion for a particle in a central field of force have already been given (eqs. 3.6-3) as

$$m\ddot{x} = F(r) \frac{x}{r},$$

$$m\ddot{y} = F(r) \frac{y}{r}.$$
(4.3-1)

Let us now vary the procedure of Sec. 3.6 by multiplying both sides of the first equation by \dot{x} dt and both sides of the second by \dot{y} dt. We thus have

$$d(\frac{1}{2}m\dot{x}^{2}) = \frac{F(r)}{r} \cdot \frac{1}{2}d(x^{2}),$$

$$d(\frac{1}{2}m\dot{y}^{2}) = \frac{F(r)}{r} \cdot \frac{1}{2}d(y^{2}).$$
(4.3-2)

Adding, and recalling that

$$x^2 + y^2 = r^2, (4.3-3)$$

and

$$\dot{x}^2 + \dot{y}^2 = v^2, \tag{4.3-4}$$

where v is the magnitude of the velocity of the particle, we obtain

$$d(\frac{1}{2}mv^2) = \frac{F(r)}{2r} d(r^2) = F(r) dr.$$
 (4.3-5)

The indefinite integral yields

$$\frac{1}{2}mv^2 - \int F(r) dr = C. (4.3-6)$$

We can now set F(r) dr = -dV(r). Then

$$\frac{1}{2}mv^2 + V(r) = C = U,$$
 (4.3-7)

where V(r) is the potential energy of the particle in the central force field; the above expresses the fact that the field is conservative, i.e., the total energy U remains constant. Incidentally it also expresses the important fact that in a central force field the magnitude of the velocity of the particle at any point depends only on its distance from the force center and is thus independent of the path or whether the particle is approaching or receding from the center of force.

Eq. $(4\cdot3-6)$ is the *energy integral*. Of course if the function F(r) happened to be a function of the velocity or the time explicitly, eq. $(4\cdot3-7)$ would not exist and the system would not be *conservative*.

To go back to

$$F(r) = -\frac{dV}{dr},\tag{4.3-8}$$

let us note that the component changes in V with respect to x and y are given by

$$\frac{\partial V}{\partial x} = \frac{dV}{dr} \frac{\partial r}{\partial x} = \frac{x}{r} \frac{dV}{dr},$$

$$\frac{\partial V}{\partial y} = \frac{dV}{dr} \frac{\partial r}{\partial y} = \frac{y}{r} \frac{dV}{dr}.$$
(4.3-9)

Substitution into (4.3-1) then yields

$$m\ddot{x} = -\frac{\partial V}{\partial x},$$

$$m\ddot{y} = -\frac{\partial V}{\partial y}.$$
(4.3-10)

But the right-hand members in $(4\cdot3-10)$ are the component forces F_x and F_y [eq. $(3\cdot6-1)$]. Hence we have as in Sec. $4\cdot1$

$$F_x = -\frac{\partial V}{\partial x}, \quad F_y = -\frac{\partial V}{\partial y}, \qquad (4.3-11)$$

and the potential energy V, considered as a function of x and y, is a function such that its negative partial derivatives with respect to x and y respectively are the component forces along the x and y axes. From this point of view V has often been referred to as the force function. We shall still continue to use the term potential energy when we take V to measure a property of the particle. Often, however, it is extremely valuable to think of V as describing the nature of the field in which the particle is moving. For this purpose we shall define V/m, i.e., the potential energy per unit mass for a particle at any particular point as the potential of the field at that point and denote it by V_P . This viewpoint may be made clearer as follows. From the definition given in Sec. 1·10, it follows that

$$F(r) dr = dW ag{4.3-12}$$

is the work done when the particle of mass m is displaced a distance dr in the direction of the force. Then

$$\int_{r_0}^{r_1} \frac{F(r)}{m} dr = \frac{V_0}{m} - \frac{V_1}{m} = V_{P_0} - V_{P_1}$$

is called the difference in potential at the points distant r_1 and r_0 from the force center. It is thus the work done in the motion of a particle of unit mass in the force field from the one place to the other, and it is interesting to see that this work is independent of the path followed by the particle, for it depends on the values of V_P at the two points only.

¹ More commonly the negative of V has been taken as the force function. This accounts for the difference in sign often encountered in various texts in the corresponding potential formulas for gravitational attraction and electrostatic attraction respectively.

² This is the definition for the case of gravitational attraction. We can also define the electrostatic potential as the potential energy per unit charge and the magnetostatic potential as the potential energy per unit magnetic pole in the case of electric and magnetic fields respectively. In general then the potential at any point in a conservative field is the potential energy per unit "quantity" of a particle at that point.

4.4. Inverse Square Field. Electron Energies in the Bohr Theory. In the case where the central field follows the inverse square law, $F(r) = -c/r^2$. Hence

$$\int F(r) dr = \frac{c}{r} + \text{const.}$$
 (4.4-1)

The total energy in such a field is therefore

$$\frac{1}{2}mv^2 - \frac{c}{r} = U,$$
 (4.4-2)

where the potential energy appears as

$$V(r) = -\frac{c}{r} {4.4-3}$$

We have already noted that the potential energy is not an absolute quantity for a given particle. In the integration (4.4-1) the constant is arbitrary and may have any value we choose to give it. It therefore follows that, in measuring potential energy, all we can really do is to give the difference in potential energy for a certain particle at two different places. If, however, in the special case of the inverse square motion, we agree to let the potential energy be zero at a very great distance from the force center, i.e., $r = \infty$, the value of the constant is zero and the potential energy is simply V = -c/r. The total energy U is thus equal to the kinetic energy at $r = \infty$. V divided by the mass of the particle also appears, as we have seen, as the potential of the field at any point distant r from the force center, i.e., the amount of work which is done on a particle of unit mass as it moves from the point in question to infinity (physically speaking a very great distance), or the work, equal to the former in amount, which is done by the particle in moving from an infinitely distant point to the point in question.

We have already seen in Sec. 3.7 that a particle in a central field of force varying inversely as the square of the distance to the force center moves in a conic section. In the particular case in which the motion is bounded the conic is an ellipse with the force center at one focus. We discussed in Secs. 3.8 and 3.9 important applications of this to planetary motion and the motion of an electron in the Bohr atomic model respectively. It is now of

interest to introduce energy considerations in this type of motion. Let us get an expression for the total energy in elliptical motion in terms of a parameter connected with the elliptical orbit. From eq. $(4\cdot4-2)$ we see that

$$U = \frac{1}{2} m v_{r_{\text{max}}}^2 - \frac{c}{r_{\text{max}}}, \qquad (4.4-4)$$

where $v_{r_{\text{max}}}$ is the velocity at aphelion. We can get this from (3.7-12) in the form

$$v_{r_{\text{max}}}^2 = \frac{k^2}{r_{\text{max}}^2}, \tag{4.4-5}$$

where k is the area constant. Now from (3.8-9) there follows

$$k^2 = \frac{ac(1 - \epsilon^2)}{m} \cdot \tag{4.4-6}$$

At the same time (3.8-20) yields

$$r_{\max} = a(1 + \epsilon). \tag{4.4-7}$$

Hence we have ultimately

$$U = \frac{c(1-\epsilon)}{2a(1+\epsilon)} - \frac{c}{a(1+\epsilon)} = -\frac{c}{2a}$$
 (4.4-8)

Since for an attractive inverse square field c is positive, this brings out again the important fact that the total energy is negative for an elliptical orbit. The formula $(4\cdot4-8)$ is attractively simple. We can apply it at once to obtain the allowed energy values for the electron orbits in the Bohr model of hydrogen.

In (3.9-14) we found the allowed or quantized values of the semi-major axis of the electron orbit in hydrogen in the form

$$a = (n_1 + n_2)^2 h^2 / 4\pi^2 m e^2$$
.

Substitution into (4.4-8) immediately yields the allowed energy values (putting $c = e^2$ in this case)

$$U_n = -\frac{2\pi^2 m e^4}{n^2 h^2}, \qquad (4.4-9)$$

where we have set $n = n_1 + n_2$, the so-called principal quantum number. The smallest value of U corresponds to n = 1 and is

called the energy of the ground state. Thus

$$U_1 = -\frac{2\pi^2 m e^4}{h^2} = -\frac{e^2}{2a_0}, \qquad (4.4-10)$$

where a_0 is the smallest possible value of a. Substitution of the relevant values of m, e, and h yields approximately

$$a_0 = 0.53 \times 10^{-8}$$
 cm.

This provides an idea of the size of the hydrogen atom in the ground state. Substitution yields

$$U_1 = 21.5 \times 10^{-12} \,\mathrm{erg}.$$

It is customary to use as the unit of energy in atomic problems the so-called electron volt, which is the energy gained by an electron in falling through a potential difference of 1 volt. The charge on the electron is 4.8×10^{-10} electrostatic units and the volt is 1/300 e.s.u. of potential. Hence the energy in question is 1.6×10^{-12} erg. Therefore

$$U_1 = 13.5$$
 electron volts.

This is termed the ionization potential of hydrogen, since it is the energy which must be given to the hydrogen atom in its ground state in order to remove the electron completely, i.e., ionize the atom.

Once the allowed energy values U_n (eq. 4.4-9) have been found for the hydrogen atom, the Bohr theory frequency postulate enables us to calculate the frequencies of the lines in the spectrum of radiation from excited and luminous hydrogen gas, as in a discharge tube. According to this postulate when a hydrogen atom makes a transition from the stationary state with energy U_{n1} to the state with energy U_{n2} , if $U_{n1} > U_{n2}$, radiation is emitted with frequency ν , where

$$\nu = \frac{U_{n1} - U_{n2}}{h} \cdot \tag{4.4-11}$$

If $U_{n1} < U_{n2}$, the atom is said to move into an *excited* state. It can then absorb radiation of the same frequency given in $(4\cdot4-11)$. One of the conspicuous successes of the Bohr theory was its ability to predict the frequencies of the lines in the so-called Balmer

series in the visible spectrum. For this, $n_2 = 2$, while $n_1 = 3, 4, 5 \cdots$. Thus

 $\nu = \frac{2\pi^2 m e^4}{h^3} \left(\frac{1}{4} - \frac{1}{n_1^2} \right) \tag{4.4-12}$

gives the observed frequencies in question with very high accuracy when $n_1 = 3, 4, 5 \cdots$. For further details on the Bohr theory the books referred to in Sec. 3.9 may be consulted.

4.5. Potential Calculations. The calculation of the potential for a field of force is often an important problem. Once it is accomplished we can write the energy equation for the motion of a

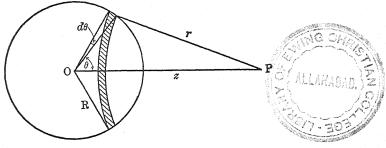


Fig. 4.2

particle in that field. Hence much attention has been paid to this problem, particularly in the case where the force field is produced by collections of mass particles. We have already noted that, according to Newton's law of universal gravitation, every mass particle attracts every other one with a force which varies directly as the product of the masses and inversely as the square of the distance between them. For a single mass particle the force field is the simple one discussed in this section; for a collection of particles the field may be expected to be in general more complicated. We shall make here a few such calculations for special cases.

Let us take first that of a thin homogeneous spherical shell of radius R. We have already computed the attractive force exerted by such a shell on an external mass particle in Sec. 2·3. At present we are more interested in the *potential* than the force. In Fig. 4·2 the shell is supposed to have its center at O. The external point is at P distant z from O. As in the previous calculation (Sec. 2·3),

we consider as the element of mass of the shell the surface element of area

 $2\pi R \sin \theta \cdot R d\theta$

obtained by drawing about P the two spheres with radii r and r+dr respectively. If m is the total mass of the shell, the mass of the element will be

$$\frac{m}{4\pi R^2} 2\pi R^2 \sin\theta \, d\theta = \frac{m}{2} \sin\theta \, d\theta, \qquad (4.5-1)$$

and the *potential* of the field produced by this element at P, which is by definition (Sec. 4.3) the potential energy of a unit mass particle placed at P, becomes,

$$V_P = -G \frac{m}{2} \frac{\sin \theta \, d\theta}{r}, \qquad (4.5-2)$$

where G is the constant of gravitation (see Sec. 2.3). In writing this we are assuming a very important fact about the potential, namely that if we have a number of mass particles the potential produced by them collectively at any point is equal to the sum of the individual potentials.

We can prove this, quite generally, as follows. Suppose we have two mass particles of mass m_1 and m_2 , and desire to calculate the potential of the field due to the two (there is no point in adding to the complexity by considering more) at the point (x, y, z) distant r_1 from mass m_1 and r_2 from mass m_2 . The force on a unit mass distant r_1 from m_1 is F_1 whose magnitude is Gm_1/r_1^2 , and similarly the force on unit mass distant r_2 from m_2 is F_2 with magnitude Gm_2/r_2^2 . Now the work done in a small displacement dr by the resultant of the two forces is

$$(\mathbf{F}_1 + \mathbf{F}_2) \cdot d\mathbf{r} = \mathbf{F}_1 \cdot d\mathbf{r} + \mathbf{F}_2 \cdot d\mathbf{r},$$

since the distributive law holds for vectors. Therefore by definition the potential V_P at the point (x, y, z) is

$$V_P = \int_{(x,y,z)}^{\infty} (\mathbf{F}_1 + \mathbf{F}_2) \cdot d\mathbf{r}$$

$$= \int_{(x,y,z)}^{\infty} \mathbf{F}_1 \cdot d\mathbf{r} + \int_{(x,y,z)}^{\infty} \mathbf{F}_2 \cdot d\mathbf{r}$$

$$= (V_P)_1 + (V_P)_2.$$

Thus the additive property follows.

We may now return to eq. (4.5-2). To obtain the expression for the potential at the point P we have merely to carry out the integration

$$V_P = -\frac{Gm}{2} \int_0^{\pi} \frac{\sin \theta \, d\theta}{r} \cdot \tag{4.5-3}$$

It will probably be most convenient to use r as the integration variable. Now

$$r^2 = R^2 + z^2 - 2Rz\cos\theta,$$

whence

$$r dr = Rz \sin \theta d\theta, \qquad (4.5-4)$$

and therefore

$$V_P = -\frac{Gm}{2} \int_{z-R}^{z+R} \frac{dr}{Rz} = \frac{-Gm}{z}$$
 (4.5-5)

The conclusion to which we are led by this result is: the *potential* of the field due to a homogeneous spherical shell at a point outside the shell is that which would be obtained if all the mass of the shell

were concentrated at the center. In Sec. 2.3 we had a similar result for the field intensity. Now as in that section we can consider a solid sphere as formed of concentric shells. Hence the result just obtained holds for the case of a solid sphere also.

It is important to find the potential due to a spherical shell at a point *inside* the shell. Consulting Fig. 4.3 and pursuing the same general plan as before, we note that the contribution to the potential at P

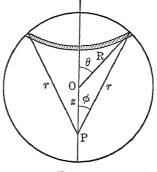


Fig. 4.3

(distant z from the center of the shell) due to the area element $2\pi R^2 \sin \theta d\theta$ is

$$-\frac{Gm}{2}\frac{dr}{zR}$$
,

recalling here eq. (4.5-4). We then have at once

$$V_P = -\frac{Gm}{2zR} \int_{R-z}^{R+z} dr = -\frac{Gm}{R}.$$
 (4.5-6)

This result is interesting and significant in that V_P is shown to be independent of the position of P within the shell and dependent only on the mass and radius of the shell. It is therefore constant everywhere inside and equal to the value at the surface. Now since the force in the direction z is $-\partial V/\partial z$, it then develops that the force on any mass particle placed inside the shell is zero; or we may say that the intensity of the gravitational field (i.e., the force

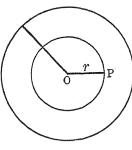


Fig. 4.4

on a unit mass) inside a homogeneous shell vanishes. These results must be modified for the case of a *solid* sphere as far as internal points are concerned.

Suppose that we have a homogeneous solid sphere of radius a and mass M, and wish to calculate V at a point inside. Consulting Fig. 4.4, let P be the point in question (distant r from O, the center), and suppose that unit mass is placed there. Now draw the sphere through P with O as center. This di-

vides the solid sphere into two parts. The part outside the sphere of radius r exerts no force on the mass particle at P, as we have just seen. The attractive force exerted on the particle by the inner sphere is G times the mass of the inner sphere divided by the square of the distance from O to P, namely r (see Sec. 2.3). Hence it is

$$-\frac{GM_{i}}{r^{2}} = -G\frac{4}{3}\pi r^{3}\frac{\rho}{r^{2}}, \qquad (4.5-7)$$

where ρ is the density of the material composing the sphere. But $\rho=M/\frac{4}{3}\pi a^2$ and hence the force on the unit mass at P is in magnitude

$$F = -\frac{GMr}{a^3} \cdot \tag{4.5-8}$$

Its direction is, of course, radial and the minus sign indicates attraction toward the center. To get the potential at P we must calculate the work required to move the unit mass from P to infinity. This may be done by direct definite integration as in (4.5-5) and (4.5-6). However, since the integration is neces-

sarily in two parts, a somewhat simpler scheme is provided by going back to (4.3-8) and writing

$$V_P = -\int F \, dr + C, \tag{4.5-9}$$

where C is a constant of integration to be evaluated from the known value of V_P at the surface of the sphere, namely -GM/a. We have (from 4.5–8)

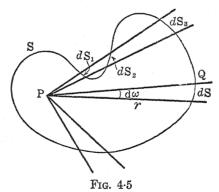
$$V_P = \frac{GMr^2}{2a^3} + C. (4.5-10)$$

But the application of the boundary condition at once yields C = -3GM/2a and gives

$$V_P = \frac{GMr^2}{2a^3} - \frac{3GM}{2a} \cdot \tag{4.5-11}$$

The calculation of the potential at points in gravitational fields due to other aggregates of mass particles will be indicated in the problems at the end of the chapter.

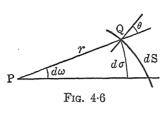
4.6. Further Considerations on the Potential. Gauss' Law and Laplace's Equation. The characterization of a conservative inverse square force field by means of the potential at any point,



or what is the same thing (as we have seen), the *potential energy* for a unit mass at that point, is of such importance in mechanics that it will be well to consider it a little more fully.

The discussion of this section will relate to space of three dimensions and be quite general. Consulting Fig. 4.5 suppose we have

a closed surface S which may or may not enclose certain mass particles. Let us consider one of these mass particles (of mass m_1)



located at the point P. Pick out an element of area dS distant r from P and draw the infinitesimal cone with dS as base and P as vertex. Examining the figure (Fig. 4.6) presenting an enlargement of Fig. 4.5, let the outward drawn normal to the surface element dS at Q make

an angle θ with PQ. Draw about P the sphere of radius PQ = r and let the area cut off by the cone on this sphere be $d\sigma$. We then have at once

$$d\sigma = dS \cos \theta. \tag{4.6-1}$$

Now the force which the mass particle m_1 exerts on a unit mass particle at Q has the magnitude

$$-\frac{Gm_1}{r^2}, \qquad (4.6-2)$$

and is directed toward P.

The component of this along the outward drawn normal is

$$-\frac{Gm_1}{r^2}\cos\theta. \tag{4.6-3}$$

Form the product of this by the area dS, whence by virtue of eq. (4.6-1) we have

$$F_N dS = \pm \frac{Gm_1}{r^2} d\sigma, \qquad (4.6-4)$$

where we have denoted the normal force component by F_N , and where the - or + sign is to be taken according as F_N is in the same direction as the normal or the contrary. Now if we construct about P as center the sphere of unit radius, and $d\omega$ is the area of the sphere cut out of the cone just mentioned, we shall have

$$d\omega = \frac{d\sigma}{r^2},$$

from the geometric properties of the cone. Hence (4.6-4) becomes

$$F_N dS = \pm Gm_1 d\omega. (4.6-5)$$

Of course, $d\omega$ is an element of solid angle about the point P. Now if every straight line from P cuts the surface in question in but one point, we can at once integrate $F_N dS$ over the whole surface and have

$$\iint F_N dS = -Gm_1 \int d\omega = -4\pi Gm_1, \qquad (4.6-6)$$

since the total solid angle about any point is 4π . This is an extremely important result, but in order to assure for it general validity and in particular justify the negative sign we must examine the case illustrated in Fig. 4.5, where some straight lines from P cut the surface in more than one point. It is to be noted, however, that the *number* of cuttings in this case (no matter how complicated the surface) must always be *odd* if P is within the surface. For example, let us take the case of *three* intersections. Here, if we pay due regard to the + and - signs in eq. (4.6–4), we have

$$F_{N_1} dS_1 = -Gm_1 d\omega,$$

 $F_{N_2} dS_2 = +Gm_1 d\omega,$
 $F_{N_3} dS_3 = -Gm_1 d\omega.$

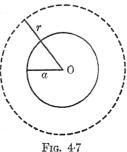
Hence the total contribution to the integral from these three will be simply $-Gm_1 d\omega$. Hence the theorem as stated mathematically in eq. (4.6-6) is generally valid. If now we proceed to apply it to the other mass particles enclosed in the surface, we have

$$\iint F_N dS = -4\pi G(m_1 + m_2 + \cdots)$$
= -4\pi GM, (4.6-7)

where M is the total mass of all the particles inside the surface. We may now state in general form the theorem we have just proved. The surface integral of the normal component of the gravitational force on a particle of unit mass (i.e., the normal component of the intensity of the gravitational field) taken over any closed surface is equal to the negative of $4\pi G$ times the total mass enclosed by the surface. This is generally known as Gauss' law of normal force. The reader may prove that it holds also for surfaces which contain no mass particles, i.e., if there are mass particles outside a closed surface, the total surface integral of the

gravitational force due to these particles is zero. This surface integral on the left-hand side of eq. (4.6-7) is sometimes called the aravitational flux.

Gauss' law is often of use in determining the gravitational intensity and potential of a field produced by a complicated



aggregate of mass particles. Here we shall note merely as an illustration the case of the homogeneous sphere which has already been discussed by other methods in Secs. 2.3 and 4.5. Consider the sphere of radius a and mass M (Fig. 4.7). Draw about the center O a sphere of radius r. From symmetry it follows that the gravitational field of force at all points of this sphere will be the same, and

will be directed normally to the spherical surface. Calling the intensity F and its normal component F_N , it follows from Gauss' law that

$$\iint F_N dS = F \iint dS = 4\pi r^2 F$$

$$= -4\pi GM. \tag{4.6-8}$$

Hence we have

$$F = \frac{-GM}{r^2} {\cdot} {(4.6-9)}$$

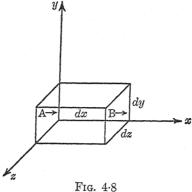
This agrees with the result of eq. (2.3-12), and from it follows, of course, that the potential at an external point distant r from the center of a solid sphere is

$$V = \frac{-GM}{r}, (4.6-10)$$

in agreement with eq. (4.5-5).

It is interesting to note that the essential content of Gauss' law can also be expressed in the form of a differential equation — and indeed one of the most important in theoretical physics. Let us suppose (Fig. 4.8) that we have an infinitesimal rectangular parallelepiped, i.e., a box of dimensions dx, dy, dz, placed for con-

venience at the origin of coordinates. Let the average gravitational field intensity over the surface A (in the yz plane) in the x



direction be F_x . The value of this quantity over B will then be (neglecting differentials of order higher than the first)

$$F_x + \frac{\partial F_x}{\partial x} dx.$$

The outward gravitational flux through the surfaces A and B then is

$$\left(F_x + \frac{\partial F_x}{\partial x} dx\right) dy dz - F_z dy dz$$

$$= \frac{\partial F_x}{\partial x} dx dy dz. \tag{4.6-11}$$

In similar fashion we find that the flux through the pair of surfaces normal to the y direction is $\frac{\partial F_y}{\partial u} dx dy dz$, while that through

the corresponding surfaces normal to the z direction is $\frac{\partial F_z}{\partial z} dx dy dz$. Hence the total normal flux out of the box is

$$\left(\frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z}\right) dx dy dz.$$

Now by Gauss' law this must be equal to $-4\pi G$ times the mass of the material in the box. The latter can be expressed as the product of the density ρ (whether the distribution is continuous or discrete) and the volume $dx\ dy\ dz$. We have, therefore:

$$\frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z} = -4\pi G\rho. \tag{4.6-12}$$

The partial differential equation of the first order $(4\cdot6-12)$ may be looked upon as Gauss' law in differential form. The quantity on the left is called the *divergence* of the gravitational field intensity **F**. Consulting Sec. $4\cdot3$ and in particular eq. $(4\cdot3-11)$, we have the general relations between **F** and V,

$$F_x = -\frac{\partial V}{\partial x}$$
, $F_y = -\frac{\partial V}{\partial y}$, $F_z = -\frac{\partial V}{\partial z}$.

It therefore follows that eq. (4.6-12) can be written equally well as a second order equation, viz.,

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} = 4\pi G\rho. \tag{4.6-13}$$

This is called *Poisson's equation* and is of great significance throughout physics, finding application in hydrodynamics and electricity and magnetism as well as in mechanics proper. When it is integrated it gives the potential V due to any distribution of mass of density ρ in the neighborhood of that distribution. In *free* space, i.e., space unoccupied by material particles, we have $\rho=0$, and the equation becomes

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} = 0, \qquad (4.6-14)$$

which is known as Laplace's equation, and is also of great importance in theoretical physics. We shall not concern ourselves with the question of the solution of equations $(4\cdot6-13)$ and $(4\cdot6-14)$ at this place, except to ask the reader to verify the statement that the potential V due to a single mass particle satisfies Laplace's equation. We recall that in this case

$$V=\frac{-Gm}{r}$$
,

where $r = \sqrt{(x - x_1)^2 + (y - y_1)^2 + (z - z_1)^2}$, the point (x_1, y_1, z_1) being the position of the particle, while (x, y, z) is the place where V is to be computed.

It will be worth while to emphasize the importance of Laplace's equation by the following statement. If we can determine a solution of it for which the potential has a definitely assigned value at every point of the boundary of a given region, the potential is uniquely determined everywhere in space. For the proof of this the student should consult a more advanced treatise.¹ Incidentally, this is a very good illustration of the power of boundary conditions in physics. The solution of Laplace's equation to satisfy given boundary conditions is in general a mathematical problem of considerable difficulty. All that we wish to emphasize here is its perfectly definite physical significance.

The reader will probably have observed that the present section seems to contain more geometry than physics. What physics is involved comes in by virtue of the *inverse square* law of gravitational attraction. So important is this law, however, that, as has been pointed out, the results of this section are of enormous significance throughout the whole of physics. For it will be recalled that this is the law which describes the attraction and repulsion of electric charges and also the mutual actions of magnetic poles. Hence the mathematical analysis of this section can at once be applied with a few changes in notation to the mathematical theory of electricity and magnetism.

PROBLEMS

- 1. A particle of mass M is supported by a vertical spring of length l and linear density ρ attached to a rigid ceiling. Find the kinetic energy of the whole system consisting of the particle and spring when the particle is pulled down a short distance and released. Assume that the vertical displacement of any point of the spring is proportional to the distance below the point of suspension. If the stiffness of the spring is k, derive the expression for the frequency of vibration taking into account the mass of the spring.
- 2. A particle of mass m moves under gravity on the parabola $y^2 = 2px$ from the point $(a, \sqrt{2pa})$ to the origin. Use the space integral of the resultant force along the path to calculate the resultant kinetic energy attained at the origin and compare with that predicted by the conservation of mechanical energy.
- 3. A simple harmonic oscillator of mass 100 grams and stiffness 10⁵ dynes/cm has a total energy of 2 joules. What is the maximum displacement of the oscillator and what is its maximum velocity? If we imagine a plane in

¹ See, for example, L. Page, Introduction to Theoretical Physics (D. Van Nostrand, N.Y., 2d ed., 1935) p. 266 ff.

which the coördinates of any point are the position x and velocity \dot{x} respectively of a particle, what kind of curve will represent the motion of the oscillator? What physical significance can be associated with the area enclosed by the curve?

- 4. A particle of mass m is subject to both a restoring force -kx (k>0) and a constant force F directed in the x direction. Find the expression for the total energy of the particle. If the latter has the value U, where will the kinetic energy of the particle vanish? Where will the potential energy vanish? Determine the displacement x of the particle as a function of the time.
- 5. Find the expressions for the kinetic energy of a particle moving in an inverse square central force field at perihelion and aphelion. Find the corresponding numerical values for the earth in its motion around the sun and for an electron moving about a nucleus with equal positive charge (hydrogen atom). (NB. Consult tables for necessary constants.)
- 6. The singly ionized helium atom consists of a single electron moving about a positively charged nucleus with mass four times that of the hydrogen nucleus and charge = 2e. Find the energy necessary to produce a doubly ionized helium atom. Between what energy states of ionized helium will a transition yield the same frequency as that of the longest wave length line in the Balmer series of hydrogen?
- 7. Compute the work required to raise a mass m from the surface of the earth to a point P, 1000 miles above the surface; also the work required to raise m from the center of the earth to the surface. What is the potential at P and at the center of the earth?
- 8. A sphere of radius a, mass M and density varying directly as the distance from the center is built up of matter brought from an infinite distance. Find the work W done throughout the process by the attraction of the matter which has already arrived on that which is brought up later, in terms of M, a, and the constant of gravitation G.
- 9. A particle moves in a central force field characterized by the potential energy $V(r) = -V_0 r e^{-\alpha r^2}$, where V_0 and α are positive constants. (a) Plot the function V(r). (b) Find the magnitude of the force as a function of r. (c) If the particle starts from rest at the point of inflection of V(r), will it reach the origin and, if so, how much velocity will it acquire? (d) How much velocity must be given the particle at the origin to remove it to infinity with zero velocity? (e) If the particle starts from rest at a distance $\ll 1/\sqrt{2\alpha}$ from the minimum of the potential energy curve what sort of motion does it perform? Discuss the characteristics of the motion.
- 10. Find the potential due to an infinitely long cylinder, of mass M per unit length and radius a, at an external point whose distance from the axis of the cylinder is r.
- 11. A homogeneous hemispherical solid has density ρ_0 and radius α . Find the expression for the gravitational potential at the center of the corresponding sphere. Calculate the components of the intensity of the gravitational field at the same point.

- 12. Derive the equation for the family of equipotential surfaces in the neighborhood of two point charges +e and -e separated by distance a.
- 13. Prove that the following are solutions of Laplace's equation and hence suitable for potential functions:

$$V = \arctan \frac{y}{x},$$

$$V = \log \frac{r+z}{r-z},$$

$$V = \frac{1}{r} \arctan \frac{y}{x},$$

$$V = \frac{z}{r^2}.$$

(These are all called spherical harmonics. Note that $r^2 = x^2 + y^2 + z^2$.)

14. Show that if V = V(r) where $r^2 = x^2 + y^2 + z^2$, Poisson's equation takes the form

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{dV}{dr}\right) = 4\pi G\rho.$$

Show that this is also equivalent to

$$\frac{1}{r}\frac{d^2}{dr^2}(rV) = 4\pi G\rho.$$

- 15. Integrate Poisson's equation to find V(r) for the case of a homogeneous spherical distribution of mass, i.e., $\rho = \text{constant}$ in Problem 14. How may the constants of integration be evaluated?
- 16. Integrate Poisson's equation to find V(r) for the case of a spherical distribution of mass in which $\rho = \rho_0 e^{-\alpha r}$, where α is a positive constant. What is the total mass of the distribution?



CHAPTER V

STATICS OF A PARTICLE

5.1. Equilibrium of a Particle. Simple Cases. In the previous chapters we have discussed the *motion* of a particle in an inertial system, and have seen that the accelerated motion of such a particle is closely associated with the idea of *force*. Now it often happens that a particle is at rest or moves in such a way that it suffers no acceleration, i.e., possesses constant velocity. The study of a particle in such a case, particularly that of rest, is usually called *statics*. It might at first be thought that the study of statics is a barren one, since the particle in this case is so restricted in its behavior. However, as soon as we are willing to admit that a particle may be subject to *two* or *more* accelerations simultaneously we see that the possibility exists for it to be at rest under those conditions, and this fact may be of great value in relating the various accelerations to which it is subject. In Chapter I we have already defined the force acting on a particle. Let us now con-

sider a mass particle on which n forces F_1, F_2, F_3, \ldots , F_n act. These may be replaced, in so far as their action on the particle is concerned, by one force R, the vector sum of all the forces. We then have for the equation of motion of the particle

$$ma = R = F_1 + F_2 + \cdots + F_n.$$
 (5·1-1)

Now if the particle is at rest or, at most, moving with constant velocity, a = 0, and we therefore have

$$R = \sum_{i=1}^{n} F_i = 0. (5.1-2)$$

 ψ_{mg} F₁, F₂, ..., F_n are then said to form a system in equilibrium. We shall call forces of this character Fig. 5-1 static forces. Let us illustrate this important idea with a few simple examples.

Consider first a particle of mass m suspended by a weightless cord as in Fig. 5·1. The particle is subjected to a gravitational force acting vertically downward and equal in magnitude to mg.

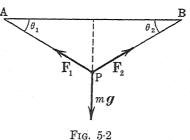
But the particle is at rest. Therefore it must be acted on by another force F_t directed upward, such that

$$\mathbf{F}_t + m\mathbf{g} = 0. \tag{5.1-3}$$

 \mathbf{F}_t is the so-called tension in the cord; a weightless spring balance or dynamometer inserted anywhere in this cord would record this force. We are not really introducing an entirely new concept of force, as might perhaps be supposed. We may still consider \mathbf{F}_t to be measured by the acceleration it could produce in a given mass, but actually the equilibrium condition $(5\cdot 1-2)$, of which

(5·1-3) is a special case, enables us more easily to compare forces of this type.

Consider another illustration. Let a particle of mass m at P (Fig. 5·2) have two weightless strings attached to it, the other ends of the strings being fastened respectively to points A and B on the same horizontal



level of a rigid ceiling. The strings being inextensible, the mass m is at rest and therefore in equilibrium. If we assume a force F_1 in string PA and a force F_2 in string PB, m will be in equilibrium provided that

$$F_1 + F_2 + F_g = 0,$$
 (5·1-4)

where \mathbf{F}_{σ} is the force of gravity, $m\mathbf{g}$. The magnitudes of \mathbf{F}_1 and \mathbf{F}_2 can be calculated by writing the equilibrium equation in scalar form. This is an important feature of the solution of problems in statics and hence we may note it particularly in connection with this simple illustration. We reason thus: the resultant \mathbf{R} of the three forces must vanish. Hence its rectangular component along any line through P must likewise vanish. Let us therefore note the equation expressing the fact that the component along the line of action of \mathbf{F}_{σ} is zero. Now the component of the resultant of several forces along any line is the sum of the components of the individual forces along this line, for a force is a vector, and this statement is true of vectors in general. We therefore have

$$F_1 \sin \theta_1 + F_2 \sin \theta_2 = mg. \qquad (5.1-5)$$

This equation alone is insufficient to determine F_1 and F_2 in terms of θ_1 , θ_2 and mg. But we can express the same condition for the component of the resultant along the line perpendicular to the vertical. This yields

$$F_1 \cos \theta_1 - F_2 \cos \theta_2 = 0.$$
 (5·1-6)

The two equations (5·1-5) and (5·1-6) are sufficient to determine F_1 and F_2 as follows,

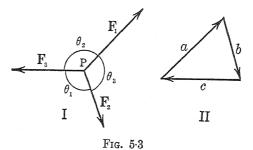
$$F_{1} = mg \cdot \frac{\cos \theta_{2}}{\sin (\theta_{1} + \theta_{2})},$$

$$F_{2} = mg \cdot \frac{\cos \theta_{1}}{\sin (\theta_{1} + \theta_{2})}.$$

$$(5.1-7)$$

Of course, other equations may be obtained by writing the above condition for any other line through P, but they will not be independent of $(5\cdot1-5)$ and $(5\cdot1-6)$.

The equilibrium of a particle under the action of three forces lying in the same plane provides an interesting theorem. Suppose



(Fig. 5·3, I) the three forces F_1 , F_2 and F_3 acting on a mass particle at P are in equilibrium. Introduce the angles θ_1 , θ_2 and θ_3 . Now let us write the conditions that the components of the resultant shall be zero in directions normal to F_3 , F_2 , F_1 respectively. Examining the figure we see that these conditions are

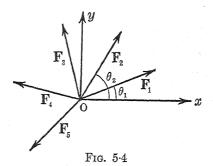
 $F_1 \sin \theta_2 = F_2 \sin \theta_1,$ $F_1 \sin \theta_3 = F_3 \sin \theta_1,$ $F_2 \sin \theta_3 = F_3 \sin \theta_2.$ These, however, may be combined at once into the two equations

$$\frac{F_1}{\sin \theta_1} = \frac{F_2}{\sin \theta_2} = \frac{F_3}{\sin \theta_3}, \tag{5.1-8}$$

which thus serve as a useful algebraic criterion for the equilibrium of the particle under the action of the forces. A more suggestive geometrical condition is at once noted when we recall that the resultant of several vectors is formed (Sec. 1.3, Fig. 1.3) by laying off at the end of the first vector a line equal in magnitude to and in the same direction as the second vector, and repeating in turn with all the vectors. The resultant is the line joining the initial point of the first vector to the end of the line last drawn. For vectors whose sum is zero the line last drawn will, of course, end at the initial point of the first vector and a closed polygon will be formed. In the case of three vectors this polygon will be a triangle (Fig. 5·3, II). Hence we have the theorem that when three forces acting on a particle are in equilibrium the force vectors must lie in a plane, and the triangle formed by drawing in the plane any three lines (not intersecting in a single point) parallel respectively to the force vectors will have the lengths of its sides proportional respectively to the magnitudes of the forces acting. This is known as the triangle of forces and is often useful in the solution of problems in statics such as the one indicated in Fig. 5.2 above. The student should solve this problem by this method, if necessary

assuming definite numerical values for mg, θ_1 , and θ_2 to fix his ideas.

5.2. General Equations of Equilibrium of a Particle. The general analytical equations for the equilibrium of a particle under the action of n coplanar forces can perhaps be most simply expressed in terms of the components along



two mutually perpendicular axes. Consulting Fig. 5.4, let the forces F_1, F_2, \ldots, F_n act on a particle at the point O. Establish in the most convenient way the rectangular axes Ox and Oy and let the forces make angles $\theta_1, \theta_2, \ldots, \theta_n$ respectively with the

positive direction of the x axis. In each case by θ we shall mean the smaller angle between the positive x direction and the direction of the force. If we denote the resultant of the forces by R and let R_x and R_y denote the components of R along the two axes respectively, we have

the two axes respectively, we have
$$R_{x} = F_{1} \cos \theta_{1} + F_{2} \cos \theta_{2} + \dots + F_{n} \cos \theta_{n} = \sum_{i=1}^{i=n} F_{i} \cos \theta_{i}$$

$$R_{y} = F_{1} \sin \theta_{1} + F_{2} \sin \theta_{2} + \dots + F_{n} \sin \theta_{n} = \sum_{i=1}^{i=n} F_{i} \sin \theta_{i}.$$
(5.2-1)

If we denote the x component of \mathbf{F}_1 by X_1 and the y component by Y_1 we may also write

$$R_x = \sum X_i, \quad R_y = \sum Y_i. \tag{5.2-2}$$

The resultant is then given in magnitude by

$$R = \sqrt{(\sum X_i)^2 + (\sum Y_i)^2},$$
 (5.2–3)

and the direction it makes with the x axis is given by ϕ , where

$$\tan \phi = \frac{R_y}{R_x} = \frac{\sum Y_i}{\sum X_i}$$
 (5.2-4)

Now if the forces are in equilibrium, R = 0 and therefore

$$\sum X_i = 0 = \sum Y_i. \tag{5.2-5}$$

Many illustrations of these conditions will be found in the problems at the end of this chapter.

Incidentally we may remark that results precisely similar to the above hold for a collection of non-coplanar forces. Here each force F_k may be considered to make angles α_k , β_k , γ_k with three mutually perpendicular coördinate axes. We then have analogously to $(5\cdot2-1)$

$$R_{x} = \sum F_{k} \cos \alpha_{k} = \sum X_{k},$$

$$R_{y} = \sum F_{k} \cos \beta_{k} = \sum Y_{k},$$

$$R_{z} = \sum F_{k} \cos \gamma_{k} = \sum Z_{k},$$

$$(5.2-6)$$

whence the resultant

$$R = \sqrt{R_x^2 + R_y^2 + R_z^2} = \sqrt{(\sum X_k)^2 + (\sum Y_k)^2 + (\sum Z_k)^2}.$$
 (5.2-7)

The condition for equilibrium is then R = 0, whence we must have

$$\sum X_k = \sum Y_k = \sum Z_k = 0, \qquad (5.2-8)$$

analogously to (5.2-5).

As a matter of fact we shall find that for most simple purposes in static problems coplanar forces are sufficient and hence we shall not press these three dimensional considerations.

5.3. A System of Particles. In the foregoing sections we have been concerned with forces acting on one particle only. We can extend our considerations to the equilibrium of several particles

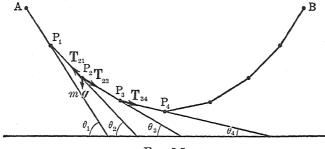


Fig. 5.5

by the use of the following idea, sometimes referred to as the principle of separate equilibrium. Suppose we have a set of particles which are connected in some way by rods or strings or otherwise act on each other. Let us imagine that these particles are likewise acted on by external forces. We can then still examine the equilibrium of any one particle apart from all the others by considering all the forces which act on it, i.e., those due to the connections as well as the external forces.

We now discuss an illustration of the above. Let a string of negligible weight be suspended from two fixed points A and B. Knotted to the string at the points P_1, P_2, \ldots, P_n are particles of equal mass m (Fig. 5.5). We are to find the angles of inclination $\theta_1, \theta_2, \ldots, \theta_n$ which the portions of the string between successive mass particles make with the horizontal. Let us first consider the equilibrium of the mass particle at P_2 . There are acting on it three forces, namely the weight mg and the two tensions T_{21} and T_{23} in the adjacent portions of the string.

The component equilibrium equations for the vertical and horizontal directions respectively are

$$T_{21} \sin \theta_2 - mg - T_{23} \sin \theta_3 = 0,$$
 $T_{21} \cos \theta_2 - T_{23} \cos \theta_3 = 0.$ (5.3-1)

Next consider P_3 and write the condition for horizontal equilibrium, namely

$$T_{34}\cos\theta_4 - T_{23}\cos\theta_3 = 0. (5.3-2)$$

It therefore follows in general (noting that $T_{12} = T_{21}$, etc.) that

$$T_{12}\cos\theta_2 = T_{23}\cos\theta_3 = T_{34}\cos\theta_4 \cdots = T,$$
 (5.3-3)

where T is some constant. Hence

$$T_{12} = \frac{T}{\cos \theta_2}$$
, $T_{23} = \frac{T}{\cos \theta_3} \cdots$
 $T_{ij} = \frac{T}{\cos \theta_j}$.

Substitution into the first of the equations (5.3-1) yields

$$\tan\theta_2=\tan\theta_3+\frac{mg}{T},$$

and the similar equations

$$\tan \theta_3 = \tan \theta_4 + \frac{mg}{T},$$

$$\cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot$$

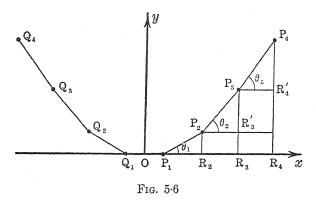
$$\tan \theta_j = \tan \theta_{j+1} + \frac{mg}{T}.$$
(5.3-4)

Let us now suppose that one of the pieces of string has a definite inclination, let us say, zero, for the sake of simplicity. That is, let $\tan \theta_{j+1} = 0$. Then

$$an heta_j = rac{mg}{T}$$
 , $an heta_{j-1} = rac{2mg}{T}$,

$$an heta_{j-2} = rac{3mg}{T}$$
,
$$an heta_1 = rac{jmg}{T} heta$$
 (5·3-5)

It thus develops in this case that the tangents of the angles of inclination are the terms of an arithmetical progression. The string we have been discussing is often called a *funicular* polygon.



It is of particular interest when the horizontal projections of the successive portions of the string are equal in length and one part is horizontal. Take the midpoint O (Fig. 5.6) of the horizontal portion Q_1P_1 as the origin of rectangular coördinates. The positions of the various mass particles to the right of O will be denoted as before by P_1, P_2, \ldots, P_j . Let the projections of these points on the x axis be R_2, R_3, \ldots, R_j respectively (note that R_1 and P_1 coincide). Let $P_2R_2 = a$, and let $Q_1P_1 = P_1R_2 = R_2R_3 = R_3R_4 = \ldots = b$. From our previous illustration we see that

$$\tan \theta_1 = \frac{a}{P_1 R_2},$$

$$\tan \theta_2 = 2 \tan \theta_1 = \frac{2a}{P_1 R_2} = \frac{P_3 R_3'}{P_2 R_3'}.$$
(5.3-6)

So we have, since $P_1R_2 = P_2R_3'$ etc.,

$$P_3 R_3' = 2a, (5.3-7)$$

$$P_4 R_4' = 3a. (5.3-8)$$

It follows that in the system of axes chosen the coördinates of P_1, P_2, P_3, \ldots are respectively $(\frac{1}{2}b, o), (\frac{3}{2}b, a), (\frac{5}{2}b, a+2a)$ etc., so that the coördinates of P_n will be $\left((2n-1)\frac{b}{2}, \frac{n(n-1)}{2}a\right)$. If we eliminate n between the expressions

$$x = (2n - 1) \frac{b}{2},$$

$$y = \frac{n(n - 1)}{2} a,$$
(5.3-9)

the result will be the equation of the curve which passes through every vertex P_n . After a little reduction we find

$$2\frac{b^2y}{a} + \frac{b^2}{4} = x^2, (5.3-10)$$

which is the equation of a parabola with Oy as its axis. The actual configuration of the string will approach more and more nearly the parabolic shape as the number of particles increases. The vertex of the parabola lies at a distance a/8 below the point O on Oy extended downwards.

The suspension bridge is an interesting illustration of equilibrium of this kind. In this case the string is replaced by a cable at intervals along which strands are placed cutting off equal intervals on the flooring of the bridge. The cable and sustaining strands may be supposed to have masses negligible compared with that of the flooring. If we assume that the masses suspended from each strand are equal to equal portions of the flooring, the conditions approximate those of the simple case just discussed. The curve of the cable will be approximately a parabola with vertex at the center of the bridge P. Let the span of the bridge be 2l and the height above P be k (see Fig. 5-7). The equation of this

parabola with P as origin will be

$$y = \frac{x^2}{4m}, \qquad (5.3-11)$$

where m is a constant. Now the slope at any point is given by

$$\frac{dy}{dx} = \frac{x}{2m} = \frac{2y}{x}. (5.3-12)$$

Hence the slope of the curve at B will be $\tan \theta_0 = 2h/l$, as we find by substituting the coördinates of B into the formula (5·3-12).

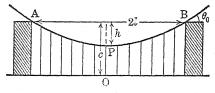


Fig. 5.7

What is the *tension* of the cable on the support at B (which of course will be equal to that at A by symmetry)? If we recall eq. (5.3-3) and those immediately following we see that the tension at the end is equal in magnitude to

$$\frac{T}{\cos\theta_0} = T \frac{\sqrt{4h^2 + l^2}}{l}, \qquad (5.3-13)$$

where the constant T is given by noting from (5.3-5) that $\tan \theta_0 = \frac{jmg}{T} = \frac{W}{2T}$, if W is the total load suspended, i.e., the weight of the

flooring of the bridge. Therefore $T=\frac{Wl}{4h}$ and the terminal tension becomes

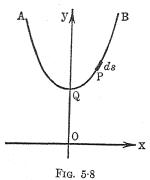
$$\frac{W}{4h}\sqrt{4h^2+l^2}.$$

Its vertical component at both B and A is of course W/2. The horizontal tension, which is everywhere the same, is, of course, T.

5.4. Equilibrium of a Flexible String. This seems a rather appropriate place to introduce the discussion of a flexible string

in equilibrium. Such a string may be considered, of course, as a very large number of small particles tied together by small weightless strings or cords, as in the example of the previous section. In the present case, however, we consider the mass to be distributed continuously along the string and further assume that the latter is perfectly flexible, i.e., can be bent in any direction without the expenditure of work.

Let us suppose such a string is suspended by two points A and B (Fig. 5-8) and acted on by gravity only. This is the simplest



case and will suffice for our illustration here as it is often encountered in nature. We desire to find the shape assumed by the string. In the first place we note that just as in the example of Sec. 5-3, the horizontal tension is the same everywhere. We shall represent distance along the string from an arbitrarily chosen origin by s. Consider an infinitesimal portion of the string of length ds at point P. Let ρ be the mass of the string per

unit length. Then if the tension is denoted by T, since the portion QP is in equilibrium the horizontal and vertical components are respectively

$$T\frac{dx}{ds}=c_1, (5.4-1)$$

$$T\frac{dy}{ds} = \rho gs + c_2, \qquad (5.4-2)$$

where c_1 and c_2 are constants. Now if we choose our origin (i.e., s = 0) at the place where the curve is horizontal, we have $c_2 = 0$, and dividing $(5\cdot4-1)$ by $(5\cdot4-2)$, there results

$$dx = \frac{c_1 dy}{\rho gs} = \sqrt{ds^2 - dy^2}, \qquad (5.4-3)$$

since $dx^2 + dy^2 = ds^2$. From this we have at once

$$dy = \frac{\rho g s \, ds}{\sqrt{c_1^2 + \rho^2 g^2 s^2}},\tag{5.4-4}$$

and the integration yields

$$y = \sqrt{s^2 + \left(\frac{c_1}{\rho g}\right)^2} + c_3. \tag{5.4-5}$$

Proceeding as above, we find x in terms of s. Thus from $(5\cdot4-3)$ and $(5\cdot4-4)$

$$dx = \frac{c_1}{\rho g} \cdot \frac{ds}{\sqrt{s^2 + \left(\frac{c_1}{\rho g}\right)^2}} \cdot \tag{5.4-6}$$

The result of the integration is

$$\frac{\rho g x}{c_1} = \log \left[s + \sqrt{s^2 + \left(\frac{c_1}{\rho g}\right)^2} \right] + c_4.$$
 (5.4-7)

The boundary condition x = 0, s = 0, gives

$$c_4 = -\log \frac{c_1}{\rho g},$$

and therefore

$$\sqrt{s^2 + \left(\frac{c_1}{\rho g}\right)^2} = \frac{c_1}{\rho g} e^{\rho g x/c_1} - s.$$
 (5.4-8)

Expanding and rearranging we have

$$s = \frac{c_1}{2\rho g} \cdot [e^{\rho g x/c_1} - e^{-\rho g x/c_1}]$$

$$= \frac{c_1}{\rho g} \sinh \frac{\rho g x}{c_1} \cdot (5.4-9)$$

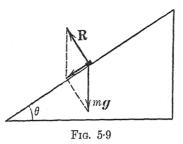
Now from eq. (5.4-5), if we let $c_3 = 0$ by choosing $y = c_1/\rho g$ at s = 0, we get at once

$$y = \frac{c_1}{\rho g} \cosh \frac{\rho g x}{c_1} \cdot \tag{5.4-10}$$

This is the equation of the famous catenary curve, as we have already noticed earlier (Sec. 2.2). The reader should plot the curve and note how it fulfills the various requirements of this section.

5.5. Equilibrium of a Particle on a Smooth Surface. In Chapter II we have already referred to the importance of the case where a

particle moves subject to certain constraints (a subject more fully investigated in Chapter VIII). We indicated there the illustration of a particle moving on an inclined plane. It becomes



of importance to notice the effect of the presence of the particle on such a surface on its equilibrium. Consider in Fig. 5.9 a particle of mass m on an inclined plane making an angle θ with the horizontal. The particle is acted on by a vertical force (gravity) of magnitude mg. This has a component perpendicular

to the plane of magnitude $mg\cos\theta$. Hence we may say that the normal thrust against the plane is $mg\cos\theta$, and by the third law of motion the plane must exert a reaction on the particle equal to $mg\cos\theta$ in magnitude and opposite in direction. We call this force R. We may then say that the particle is acted on by two forces, one that of gravity and the other the reaction of the plane. Strictly speaking, of course, this second force is merely an expression of the fact that the particle must remain on the plane. Thus if the particle were otherwise free in space but subject to the two forces mg vertically and $R=mg\cos\theta$ making an angle θ with the vertical, the resultant force would be precisely what it is with the particle constrained to move on the plane under gravity. This resultant, as is clear from the figure, acts in the direction down the plane and is equal in magnitude to $mg\sin\theta$.

We are thus led to consider that whenever a particle is constrained to rest on a surface or a curve, in addition to the other forces acting on it there will always be a reaction of the surface or curve. An important question arises as to the direction of the reaction. In the above illustration we have taken it as normal to the surface. This will always be the case when the surface is perfectly smooth, for by definition a smooth surface is one that can exert no force in a direction parallel to itself.

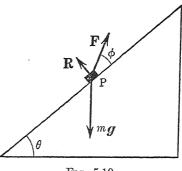
We now wish to investigate the equilibrium of a particle on such a surface. From our general conditions (Sec. 5.2) we can see at once that equilibrium will result if the reaction of the surface is equal in magnitude but opposite in direction to the resultant of all the other forces acting on the particle. We shall con-

sider here only the case where the forces lie in a plane. As an illustration consider Fig. 5·10, where the particle is in equilibrium

on an inclined plane of angle θ under gravity and a force F acting at an angle ϕ with the plane. We apply the usual conditions, taking for convenience our components along and normal to the plane. The condition along the plane is

$$F\cos\phi - mg\sin\theta = 0, \quad (5.5-1)$$

while normal to the plane we have



$$R + F \sin \phi - mg \cos \theta = 0. \qquad (5.5-2)$$

From (5.5-1) there results

$$F = mg \frac{\sin \theta}{\cos \phi}, \qquad (5.5-3)$$

while from (5.5-2) there then follows

$$R = mg \left[\cos \theta - \frac{\sin \theta \cdot \sin \phi}{\cos \phi} \right] = \frac{mg}{\cos \phi} \cos (\theta + \phi), \quad (5.5-4)$$

for the required reaction force. This procedure is perfectly general. More complicated examples will be found among those at the end of the chapter.

5.6. Equilibrium of a Particle on a Rough Surface. Static Friction. The perfectly smooth surface discussed in the preceding section is a highly idealized conception. All surfaces encountered in nature are actually rough to a greater or less degree and because of their inequalities are able to exert on particles resting on them forces which act parallel to the surface, thus resisting sliding motion. The force which is operative in this case is called friction. It is distinguished from the adhesion which bodies have for each other by the fact that while the latter is independent of the force thrusting the two surfaces together, the former is dependent on the normal thrust. This statement is one of the so-called experi-

mental laws of friction. The ratio of the force of friction to the total normal thrust for surfaces *just about to slip* over each other is a quantity which remains constant for considerable variations in the thrust. We are then led to define

$$\mu = \frac{F}{N} = \frac{\text{magnitude of maximum frictional force}}{\text{magnitude of normal thrust}}, \quad (5\cdot6-1)$$

as the coefficient of friction. It should be emphasized that the constancy of this quantity is only approximate, for μ varies greatly as N increases to a point where the surface is deformed under its action. The second law of friction states that for a given normal thrust the force of friction is approximately independent of the area of the surface of contact. Obviously this also has its limitations. One can hardly expect to apply it if the surface is a pin point, for example. The coefficient of friction depends, of course, on the nature of both surfaces in contact and varies greatly with variations in these. The following table gives some illustrations of the approximate values of μ for various cases.

COEFFICIENT OF FRICTION (From Smithsonian Tables) All values are approximate only

Materials	Coefficient of friction
Wood on wood, dry	.2550
Wood on wood, soapy	20
Metals on oak, dry	.5060
Metals on oak, wet	.2426
Metals on oak, soapy	.20
Metals on elm, dry	.2025
Hemp on oak, dry	.53
nemp on oak, wet	.33
Learner on oak	.2738
Leather on metals, dry	.56
Leather on metals, wet	.36
Leatner on metals, greasy	.23
Leather on metals, onv	.15
Metals on metals, dry	.1520
Metals on metals, wet.	.1520
Smooth surfaces occasionally greased.	.0708
Smooth surfaces, best results.	
Steel on agate, dry	.03036 .20
Steel on agate, oiled	
Iron on stone.	.107
Wood on stone.	.3070
Earth on earth	about .40
Earth on earth, damp clay.	. 25-1 . 00
Earth on earth, wet clay	1.00
The contract of the contract o	.31

As one might also expect, the coefficient μ depends on the character of the motion of one surface on the other. If the beginning motion is rolling, for example, rather than sliding we should not anticipate the same state of affairs in both cases. Moreover after motion has once been established, the coefficient of friction may be expected to be different. It is actually found that friction in sliding motion is less than static friction. It is the latter with which we are concerned in the present section.

Moreover we ought to emphasize the fact that the magnitude of the force of friction is variable. Thus if a heavy particle rests on a rough horizontal table and the only force acting on it is its weight, the force of friction is zero. If a horizontal force is applied to the particle, the force of friction will be equal to this force as long as no slipping takes place. It is *only* when slipping is about to take place that $F = \mu N$. Hence we may say in general that F can vary all the way from zero to μN .

This is well brought out by the equilibrium of a particle on a rough inclined plane (cf. Fig. 5·11). In addition to the weight

mg the particle at P is acted on by the reaction force R_1 with magnitude mg cos θ and the frictional force F, which acts along the plane in such a direction as to oppose the motion which would result from gravity alone. We may think of the forces R_1 and Fas combining to form a resultant reaction R, which if equilibrium

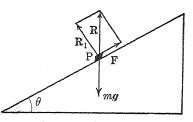


Fig. 5.11

is to be maintained must be equal in magnitude to mg and be oppositely directed to the weight. For equilibrium $F=mg\sin\theta$ always, but it will not reach its maximum possible value $\mu R_1=\mu mg\cos\theta$ until slipping is just about to take place. This will evidently happen when the angle θ reaches the value χ where

$$mg \sin \chi = \mu mg \cos \chi$$
,

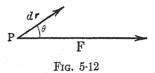
or

$$\tan \chi = \mu. \tag{5.6-2}$$

The angle χ is sometimes called the *angle of repose*. Its measurement provides a method for the determination of the coefficient of friction.

We may again emphasize the fact that frictional force is different in nature from the other forces so far discussed. Its essential nature is resistance; in fact we may call it passive resistance. It always acts in such a direction as to resist motion. Moreover, we may stress the fact that it is not a conservative force in the sense of Chapter III. There is no potential associated with it. We shall encounter similar forces in hydrodynamics and elsewhere.

5.7. The Principle of Virtual Work. An interesting interpretation of the equilibrium of a particle may be obtained by a consideration of the idea of work developed in the previous chapter. Thus consider a particle at the point P (Fig. 5.12), and imagine that it is



acted on by certain forces. Let one of these be the force **F**. We shall suppose that we do not know actually how the particle moves under the action of the forces; but imagine that an arbitrary possible small displace-

ment of the particle be denoted by $d\mathbf{r}$. This displacement, it must be emphasized, need not be one that actually ever takes place. It is sufficient that we can imagine its taking place without violating any constraints to which the particle may be subjected. For this reason it is denoted by the term virtual displacement.

The work that is done during this virtual displacement in the case considered is clearly

$$\mathbf{F} \cdot d\mathbf{r} = F \, dr \cos \theta. \tag{5.7-1}$$

This work is called the virtual work.

Let us consider a number of forces F_1, F_2, \ldots, F_n impressed on a particle at P. We shall now say that if and only if the total virtual work performed by these n forces during any virtual displacement of the particle is zero, the particle is in equilibrium under the action of the forces. This is the celebrated principle of virtual work first enunciated by John Bernoulli in 1717. In many texts it is derived from the conditions for static equilibrium already laid down. However, we shall assume it as a general principle independent of other considerations and show that the other equilibrium conditions follow from it. Consulting Fig. 5·13, let the arbitrary dis-

placement be $d\mathbf{r}$. Then the condition of equilibrium above given is expressed as

$$\mathbf{F}_1 \cdot d\mathbf{r} + \mathbf{F}_2 \cdot d\mathbf{r} + \dots + \mathbf{F}_n \cdot d\mathbf{r} = 0. \tag{5.7-2}$$

For the total virtual work is clearly the sum of the individual virtual works. But this is

$$(\mathbf{F}_1 + \mathbf{F}_2 + \cdots + \mathbf{F}_n) \cdot d\mathbf{r} = 0,$$
 (5.7-3)

since vectors obey the distributive law. But

$$\mathbf{F}_1 + \cdots + \mathbf{F}_n = \mathbf{R}, \tag{5.7-4}$$

where R is the *resultant* of the *n* forces acting on the particle. Now since $d\mathbf{r}$ is a virtual displacement, it is perfectly *arbitrary*, and the only way for $\mathbf{R} \cdot d\mathbf{r}$ to be zero for all $d\mathbf{r}$ is to have

$$R = 0.$$
 (5.7–5)

This, however, is the general condition for the equilibrium of static forces, which thus follows from the general principle of virtual work. We can also get the corresponding conditions in terms of the rectangular compo-

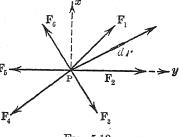


Fig. 5.13

nents. For to return to Fig. 5·13, let us establish the axes Ox and Oy at the point P. For convenience it will be assumed that the forces are coplanar, though this is of course not a necessary restriction. Then we get

$$\mathbf{F}_{1} = \mathbf{i}F_{1x} + \mathbf{j}F_{1y},
\mathbf{F}_{2} = \mathbf{i}F_{2x} + \mathbf{j}F_{2y},
\vdots
\vdots
\vdots
\mathbf{F}_{n} = \mathbf{i}F_{nx} + \mathbf{j}F_{ny},$$
(5.7-6)

while $d\mathbf{r} = \mathbf{i} dx + \mathbf{j} dy$, to use the customary vector notation (Sec. 3-1, etc.). Eq. (5-7-2) now becomes

$$(F_{1x} + F_{2x} + \dots + F_{nx}) dx + (F_{1y} + F_{2y} + \dots + F_{ny}) dy = 0.$$
 (5.7-7)

But dx and dy are of course arbitrary. Hence

$$F_{1x} + F_{2x} + \dots + F_{nx} = \sum F_{ix} = 0,$$

$$F_{1y} + F_{2y} + \dots + F_{ny} = \sum F_{iy} = 0,$$
(5.7-8)

which are the conditions for equilibrium referred to in Sec. 5.2,

and are, of course, equivalent to (5.7-5).

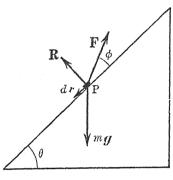


Fig. 5.14

An interesting special case is that in which the virtual displacement is so chosen as to be at right angles to one of the forces in question. This often simplifies the application of the principle, since the resulting virtual work for this particular force is zero. Thus suppose (see Fig. 5·14) that we have a particle of mass m in equilibrium on a smooth inclined plane making an angle θ with

the horizontal, under the action of a force F making an angle ϕ with the plane. If we wish to discuss the equilibrium of the

particle without reference to the reaction force \mathbf{R} we need only consider as a virtual displacement the displacement $d\mathbf{r}$ along the plane. Then $\mathbf{R} \cdot d\mathbf{r} = 0$ and the principle of virtual work says that

$$mg \sin \theta \, dr - F \cos \phi \, dr = 0,$$
 or

$$F\cos\phi = mg\sin\theta$$
, (5.7-9)

since $dr \neq 0$. The above equation has already been previously obtained by the usual method [eq. (5.5-3)].

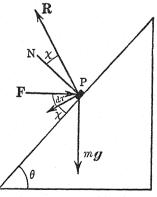


Fig. 5.15

A somewhat more significant illustration of the application of the principle is provided by the case of a particle on a rough plane. Suppose (see Fig. 5·15) that the particle P of mass m when acted

on by the horizontal force F is just on the point of slipping down the plane. In this case the reaction R of the plane makes with the normal the angle of repose χ . Pursuing the method just mentioned we consider as the virtual displacement, $d\mathbf{r}$ perpendicular to R. Then $\mathbf{R} \cdot d\mathbf{r} = 0$, as before, and the application of the principle yields

$$mg\sin(\theta - \chi) dr - F\cos(\theta - \chi) dr = 0, \quad (5.7-10)$$

or

$$F = mg \tan (\theta - \chi). \tag{5.7-11}$$

5.8. Virtual Work — Two or More Particles. Rather more interest attaches to the principle of virtual work when applied to a system of two or more connected particles. The question arises: How shall the principle be stated in this case? Suppose we label the particles $1, 2, 3, \ldots n$ and imagine that forces $F_1, F_2, F_3, \ldots, F_n$ are impressed on them respectively (i.e. F_1 on particle 1, etc.). Suppose that while acted on by these forces the particles undergo virtual displacements $d\mathbf{r}_1, d\mathbf{r}_2, \ldots, d\mathbf{r}_n$. The equation of the total virtual work to zero gives us

$$\mathbf{F}_1 \cdot d\mathbf{r}_1 + \mathbf{F}_2 \cdot d\mathbf{r}_2 + \cdots + \mathbf{F}_n \cdot d\mathbf{r}_n = 0. \tag{5.8-1}$$

It is to be remembered that the displacements must satisfy the connections and constraints of the system and in general will not be completely arbitrary, i.e., there will exist relations among them. Our problem in any particular case will be to write (5·8–1) for the given impressed forces and then to ascertain by an examination of the geometry of the situation what the relations are between the virtual displacements. This will enable us to reduce them to a set of completely arbitrary displacements and hence find the equilibrium conditions the forces must satisfy.

An illustration will make this clear. Let us consider the equilibrium of two mass particles m_1 and m_2 at the two ends of an inextensible string which slides freely over a smooth peg. We shall suppose that in addition to the impressed forces of gravity m_1 g and m_2 g there are also the extra impressed forces F_1 and F_2 as indicated in Fig. 5·16. Applying the virtual displacements dr_1 and dr_2 , the general relation (5·8-1) becomes

$$(\mathbf{F}_1 + m_1 \mathbf{g}) \cdot d\mathbf{r}_1 + (\mathbf{F}_2 + m_2 \mathbf{g}) \cdot d\mathbf{r}_2 = 0.$$
 (5.8-2)

The displacements may be taken in any way compatible with the constraints. We find it most convenient to take them vertically

 m_1 m_1 m_2 m_2 m_2 m_2

Fig. 5.16

downward. Expanding the dot products then yields

$$(F_1 + m_1 g) dr_1 + (F_2 + m_2 g) dr_2 = 0.$$
 (5.8-3)

But since the string is inextensible, dr_1 and dr_2 are not completely arbitrary. Rather

$$dr_1 + dr_2 = 0. (5.8-4)$$

The combination of (5.8-3) and (5.8-4) then yields

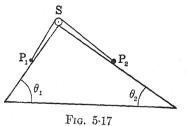
$$F_1 + m_1 g = F_2 + m_2 g, \qquad (5.8-5)$$

which solves the problem. Note that, although the conventional equilibrium treatment of this problem is forced to introduce the tension in

the string, the latter is not an impressed force and hence does not enter into the expression of the principle of virtual work.

Another illustration of similar nature is provided by two particles P_1 and P_2 of masses m_1 and m_2 respectively tied to the ends of

an inextensible string wrapped around a smooth peg S at the vertex of two perfectly smooth inclined planes of angles θ_1 and θ_2 respectively (Fig. 5·17). The impressed forces are the weights $m_1\mathbf{g}$ and $m_2\mathbf{g}$. It is unnecessary to consider the tension in the string or the reaction of the



planes since they are constraining and not impressed forces. Taking virtual displacements $d\mathbf{r}_1$ and $d\mathbf{r}_2$ down the plane in each case, we write (5.8-1) in the form

$$m_1\mathbf{g} \cdot d\mathbf{r}_1 + m_2\mathbf{g} \cdot d\mathbf{r}_2 = 0. \tag{5.8-6}$$

The expansion of the dot product gives us

$$m_1 g \sin \theta_1 dr_1 + m_2 g \sin \theta_2 dr_2 = 0,$$
 (5.8-7)

and the inextensibility of the string,

$$dr_1 + dr_2 = 0. (5.8-8)$$

Hence we are led ultimately to

$$m_1g\sin\theta_1 = m_2g\sin\theta_2. \tag{5.8-9}$$

Reverting to (5.8-1), let us express it in Cartesian form by writing for the pth particle

$$d\mathbf{r}_{p} = \mathbf{i} dx_{p} + \mathbf{j} dy_{p} + \mathbf{k} dz_{p},$$

$$\mathbf{F}_{p} = \mathbf{i} F_{xp} + \mathbf{i} F_{yp} + \mathbf{k} F_{zp}.$$
(5.8-10)

The principle of virtual work may now be expressed in the form

$$\sum_{p=1}^{n} (F_{xp} dx_p + F_{yp} dy_p + F_{zp} dz_p) = 0 (5.8-11)$$

Let us suppose that the forces are conservative (Sec. 4.1) so that there exists a potential energy function $V_p(x_p, y_p, z_p)$ such that

$$F_{xp} = -\frac{\partial V_p}{\partial x_p}$$
, etc. (5.8-12)

Then

$$\sum_{p=1}^{n} V_p = V {(5.8-13)}$$

may be considered the total potential energy of the collection of particles. The principle of virtual work then appears in the form

$$\sum_{p=1}^{n} dV_p = 0, (5.8-14)$$

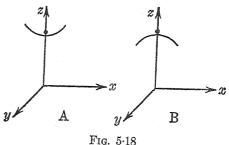
or for the virtual displacements considered the change in the total potential energy is zero. This signifies that for the equilibrium state the potential energy must be either a maximum or a minimum, for it is only in such a case that the variation due to a small displacement can vanish. In the former case the equilibrium is said to be unstable, while in the latter case it is stable. Perhaps the simplest illustration is that of a single particle acted on by gravity

¹ Mathematically this is an over simplification. Strictly speaking, all we can say is that the potential energy has a stationary value. It might for example be a mini-max. Cf. W. V. Houston, *Principles of Mathematical Physics* (N.Y., McGraw-Hill, 2d ed., 1948) p. 89 ff. For most mechanical problems it is sufficient to confine our attention to maxima and minima.

and resting on a smooth surface. Taking the z axis vertical, the potential energy to an additive constant is

$$V = mgz, (5.8-15)$$

where z is measured positively upward. Now suppose the particle is in equilibrium on a surface concave upward. The total force must act normally to the surface and hence z must be a minimum and therefore V also. This corresponds then to stable equilibrium (A in Fig. 5·18). Any finite displacement of the particle on the



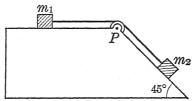
surface will increase the potential energy. On the other hand, if the particle is in equilibrium on a surface concave downward (B, Fig. 5.18), z must be a maximum and therefore V must be a maximum. The equilibrium is unstable. Any finite displacement of the particle leads to a *decrease* in potential energy.

This significant connection between equilibrium and potential energy affords another illustration of the utility of the principle of virtual work.

PROBLEMS

- 1. If a string of length l has a line density ρ and we consider its total weight concentrated at its central point, with how much tension must it be stretched between two fixed points in a horizontal line in order to make it assume a horizontal direction? Discuss the physical significance of the answer.
- 2. The magnitude of the resultant of two forces F_1 and F_2 is 14 lb wt. The force F_1 has a magnitude of 12 lb wt., and the direction of F_2 is inclined to the resultant at an angle of 30°. Find F_2 and the angle between F_1 and F_2 .
- 3. A particle of mass m is attached to one end of a string and the other end is fixed. Find the force which must be applied to the particle perpendicular to the string in order to hold the string at an angle θ to the vertical. Find also the tension in the string.

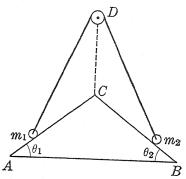
- 4. A mass of 10 kilograms is supported at rest on a smooth inclined plane by a horizontal force of 4 kilograms wt. Find the inclination of the plane. If the plane were rough with coefficient of friction between body and plane of 0.2 and if the angle of inclination were that just determined, what horizontal force would be required to support the body at rest?
- 5. In the figure $m_1 = 1$ kg and $m_2 = 3$ kg and P is a smooth peg. The masses are connected by an inextensible string. The mass m_1 rests on a horizontal plane, with coefficient of friction between block and plane = 0.2. The mass m_2 rests on a plane making the angle 45° with the horizontal. What must the coefficient of friction be between m_2 and its plane in order to maintain equilibrium?



- 6. Two equal magnets of pole strength μ and length l are placed parallel to each other at distance a. If like poles are adjacent, calculate the position or positions where a single pole would be in equilibrium under the action of the magnets (i.e., the so-called *neutral* point). Solve the same problem when the adjacent poles are unlike. (Hint: Recall Coulomb's law of force between poles, $F = \mu \mu'/r^2$.)
- 7. Given a magnet of length l and pole strength μ . Derive the equation of a line of force of the magnetic field surrounding the magnet, that is, a curve whose tangent at any point has the direction of the resultant force on a single pole placed at this point.
- 8. Derive the equation of a line of force in the field of two particles charged with equal amounts of positive electricity and separated by the distance a.
- 9. Two small metal spheres with a mass of 1 gram each are suspended from a single point by threads 20 cm long. The balls are equally charged with 20 electrostatic units of electricity. (The electrostatic unit is one such that it attracts an equal and opposite charge at a distance of one cm with a force of one dyne.) Find the distance to which the charged spheres will separate for equilibrium.
- 10. Two charged particles with charges +10 electrostatic units and -20 electrostatic units respectively interact with each other. What is the total potential energy of the system? Apply the principle of virtual work and find the conditions under which the system may be in equilibrium with each particle at rest.
- 11. If the small spheres in Problem 9 have masses 1 gram and 2 grams respectively, are charged with 10 and 20 electrostatic units respectively, and are suspended from a single point by threads 20 cm long, find the distance to

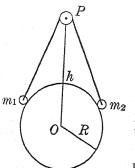
which the spheres will separate for equilibrium. Also find the angles which the two threads make with the vertical and the tensions in them.

12. Given a double inclined perfectly smooth plane ABC with angles θ_1 and θ_2 as indicated in the accompanying figure. A smooth peg is placed vertically above C so that CD = h. An inextensible string with masses m_1 and m_2 at the



two ends is passed over the peg and the masses are allowed to move on the planes. Find the position of equilibrium by the principle of virtual work.

13. Given a circle of radius R in a vertical plane. A smooth peg is placed at point P directly above the center of the circle so that OP = h. (See figure.)



A string of length l passes over the pulley and has at its extremities the two masses m_1 and m_2 . Assuming that the contact of the masses and the circle is smooth, calculate the position of the masses for equilibrium. Use the principle of virtual work.

14. A flexible cord is suspended from two points on the same horizontal line. If the cord hangs in the form of a parabola, what must be the law of variation of the mass per unit length?

15. A flexible uniform cord of length l is suspended from two fixed points in the same horizontal line. Find the distance between the two

fixed points so that each terminal tension is three times the tension at the lowest point.

16. The entire weight supported by a suspension bridge is 8 tons. The total length is 250 feet and the height is 20 feet. Find the tension on the points of support and the horizontal tension at the lowest point.

17. A telegraph line is constructed of copper wire weighing 0.079 pound per foot. The distance between the posts is 150 feet. If the tension at the posts is 200 lb wt., find the sag of the wire in the middle.

CHAPTER VI

MOTION OF A SYSTEM OF PARTICLES

6.1. Equations of Motion. Center of Mass. Conservation of Momentum. In our discussion so far we have been concerned for the most part with the motion of a single particle. Actual bodies, however, must be thought of as composed of a multitude of particles and, in order to study their behavior, it is now necessary to consider in some detail the equations governing the motion of a collection of particles. We shall approach the problem in a rather general way and try to establish some general results which will be independent of the forces of interaction holding the particles together.

Suppose we have a system of n particles of masses m_1, m_2, \ldots, m_n , respectively, and acted on by the forces F_1, F_2, \ldots, F_n , respectively. Each force will be the sum of an external force and all the interaction forces between the particle in question and all the other particles. Thus for the jth particle

$$\mathbf{F}_{j} = \mathbf{F}_{je} + \sum_{k(\neq j)=1}^{n} \mathbf{F}_{jk},$$
 (6·1-1)

where \mathbf{F}_{je} is the external force on the *j*th particle, and \mathbf{F}_{jk} is the force on the *j*th particle due to the *k*th particle. From eq. (1·7-4) the equation of motion of the *j*th particle in vector form is then

$$m_j \ddot{\mathbf{r}}_j = \mathbf{F}_{je} + \sum_{k(\neq j)=1}^n \mathbf{F}_{jk},$$
 (6·1-2)

where \ddot{r}_j is the position vector of the jth particle in the chosen reference system. The summation notation in $(6\cdot 1-1)$ and $(6\cdot 1-2)$ merely expresses the fact that, in the sum, as k goes from 1 to n the term corresponding to k=j is omitted, as it is meaningless, since the jth particle does not interact with itself. Let us now sum each equation over all the particles. We obtain formally

$$\sum_{j=1}^{n} m_{j} \ddot{\mathbf{r}}_{j} = \sum_{j=1}^{n} \mathbf{F}_{je} + \sum_{j=1}^{n} \sum_{k(\neq j)=1}^{n} \mathbf{F}_{jk}.$$
 (6·1-3)

The double summation notation on the right appears a trifle

complicated but should cause no difficulty. It means that after summing \mathbf{F}_{jk} over k from 1 to n (leaving out k=j) we sum the resulting sum over j as the latter runs from 1 to n. Thus, writing it out in full,

$$\sum_{j=1}^{n} \sum_{k(\neq j)=1}^{n} \mathbf{F}_{jk} = \sum_{j=1}^{n} (\mathbf{F}_{j1} + \mathbf{F}_{j2} + \dots + \mathbf{F}_{jn})$$

$$= \mathbf{F}_{12} + \mathbf{F}_{13} + \dots + \mathbf{F}_{1n}$$

$$+ \mathbf{F}_{21} + \mathbf{F}_{23} + \dots + \mathbf{F}_{2n}$$

$$+ \dots + \mathbf{F}_{n1} + \mathbf{F}_{n2} + \dots + \mathbf{F}_{n,n-1}.$$

Now in the above expansion, for every term of the form \mathbf{F}_{jk} there is a term of the form \mathbf{F}_{kj} . But \mathbf{F}_{jk} is the force on the *j*th particle due to the *k*th particle, and \mathbf{F}_{kj} is the force on the *k*th particle due to the *j*th particle. From Newton's third law (the law of action and reaction) we must have

$$\mathbf{F}_{jk} = -\mathbf{F}_{kj}. \tag{6.1-4}$$

It follows that the double sum in $(6\cdot1-3)$ vanishes identically, and the equation can be written in the following form

$$\sum m_j \ddot{\mathbf{r}}_j = \sum \mathbf{F}_{je}, \tag{6.1-5}$$

or alternatively

$$\frac{d^2}{dt^2} \left(\sum m_j \mathbf{r}_j \right) = \sum \mathbf{F}_{je}. \tag{6.1-6}$$

This conveys the strong suggestion that we introduce a point whose position vector $\ddot{\mathbf{r}}$ is given by

$$\tilde{\mathbf{r}} = \frac{\sum m_i \mathbf{r}_i}{\sum m_i} \cdot \tag{6.1-7}$$

The motion of this point will then be given by the equation

$$m\ddot{\tilde{\mathbf{r}}} = \sum \mathbf{F}_{je}, \qquad (6.1-8)$$

where $m = \sum m_j$ = the total mass of the collection. The point with position vector $\bar{\mathbf{r}}$ moves as if all the mass of the collection were concentrated there and all the external forces were acting there. We call it the *center of mass* of the collection of particles. It is an important property of the system, since if we can follow its motion

we have at least some idea where in space the system is, even though we may be unable to follow the motion of every individual particle. Clearly (6·1-8) says that in the special case in which the sum of the external forces vanishes, the center of mass moves in a straight line with constant velocity (rest being, of course, a special case).

We can give another interpretation of eq. $(6\cdot1-6)$, since it may also be written in the form

$$\frac{d}{dt} \left(\sum m_j \dot{\mathbf{r}}_j \right) = \sum \mathbf{F}_{je}. \tag{6.1-9}$$

Now $\sum m_j t_j$ is the resultant momentum of the collection of particles, whose time rate of change is therefore equal to the resultant external force acting on the system. Once more note the important special case in which $\sum \mathbf{F}_{je} = 0$. Then the resultant momentum remains unchanged. This is the theorem of the conservation of momentum for a system of particles subject only to their own mutual interaction forces. A simple illustration is provided by the collision of two billiard balls. This is sufficiently interesting to warrant a section by itself.

6.2. Conservation of Momentum in Collisions. We imagine two hard solid spheres of masses m_1 and m_2 , respectively, moving originally with constant velocities \mathbf{u}_1 and \mathbf{u}_2 in the same straight line; we suppose that at some instant they collide squarely, with the result that after collision the spheres again separate and their velocities become \mathbf{v}_1 and \mathbf{v}_2 , respectively. Though the spheres are themselves really aggregates of particles (i.e., approximately rigid bodies), since they collide squarely we may safely treat them as single particles. Since the collision involves only interaction forces, the principle of conservation of momentum applies and we can write

$$m_1\mathbf{u}_1 + m_2\mathbf{u}_2 = m_1\mathbf{v}_1 + m_2\mathbf{v}_2.$$
 (6.2-1)

This equation does not enable us to predict the after-collision velocities from the original velocities. Consequently, the full solution of the problem demands further assumptions. Let us first write $\mathbf{u}_1 = u_1 \delta$, $\mathbf{u}_2 = u_2 \delta$, $\mathbf{v}_1 = v_1 \delta$, $\mathbf{v}_2 = v_2 \delta$, where δ is the unit vector along the straight line in which the motion takes place.

144 MOTION OF A SYSTEM OF PARTICLES

We can then replace the vector equation $(6\cdot2-1)$ by the scalar equation

$$m_1u_1 + m_2u_2 = m_1v_1 + m_2v_2. (6.2-2)$$

Now Newton, who experimented with colliding spheres, found that their behavior satisfies the further equation

$$v_1 - v_2 = -e(u_1 - u_2),$$
 (6.2-3)

where e is a positive proper fraction, called the *coefficient of restitution*. Its value depends on the material composing the spheres. For highly elastic substances, i.e., those which, though deformed by the collision, regain their original size and shape immediately thereafter, e is very close to unity, while for plastic substances in which collision leads to a more or less permanent deformation, e may be close to zero.

Let us use $(6\cdot2-2)$ and $(6\cdot2-3)$ to calculate the after-collision velocity magnitudes v_1 and v_2 in terms of the original values u_1 and u_2 . The algebra leads to

$$v_{1} = \frac{(m_{1} - m_{2}e)u_{1} + m_{2}(1 + e)u_{2}}{m_{1} + m_{2}},$$

$$v_{2} = \frac{m_{1}(1 + e)u_{1} + (m_{2} - m_{1}e)u_{2}}{m_{1} + m_{2}}.$$
(6·2-4)

The special case e = 1 is of interest. For this

$$v_{1} = \frac{(m_{1} - m_{2})u_{1} + 2m_{2}u_{2}}{m_{1} + m_{2}},$$

$$v_{2} = \frac{2m_{1}u_{1} + (m_{2} - m_{1})u_{2}}{m_{1} + m_{2}}.$$
(6.2-5)

If in addition $m_1 = m_2$, these equations become

$$v_1 = u_2, \quad v_2 = u_1, \tag{6.2-6}$$

or the two spheres merely exchange their velocities on collision.

On the other hand, suppose e = 0 (plastic impact). In this case (6·2-4) yields

$$v_1 = \frac{m_1 u_1 + m_2 u_2}{m_1 + m_2} = v_2, \tag{6.2-7}$$

which means that after collision the two spheres move off together. This is well illustrated in the impact of two putty balls.

Another important special case of collisions is that in which the mass of one particle is very much greater than that of the other. Let us suppose that $m_2 \gg m_1$ and that $u_2 = 0$. Then eqs. $(6\cdot 2-4)$ yield effectively

$$v_1 = -eu_1, \quad v_2 = 0.$$
 (6.2–8)

A practical application is the rebound of a ball of mass m_1 from a horizontal table of mass very much larger than m_1 . If the ball is dropped from height h_1 , we have

$$v_1 = -e\sqrt{2gh_1}, \quad v_2 = 0.$$

The height of the rebound of m_1 is

$$h_2 = v_1^2 / 2g = h_1 e^2. (6.2-9)$$

This provides a method of measuring the coefficient of restitution between two substances.

Further investigation of collision phenomena demands the introduction of energy considerations which we shall postpone to Sec. 6.5.

6.3. The Two-particle Problem with Gravitational Attraction. The general results set forth in Sec. 6.1 receive another interesting illustration in the special case of a system of two particles. This has immediate practical application to celestial bodies which though large are very far apart and hence may be considered as particles attracting each other in accordance with the Newtonian law of gravitation. We therefore discuss the case in which two particles attract each other with a force varying inversely as the square of their distance apart and in which no other forces act. This is a generalization of the inverse square central force field problem treated in Sec. 3.7.

If the masses of the particles are m_1 and m_2 , respectively, and their position vectors in the chosen reference system are \mathbf{r}_1 and \mathbf{r}_2 , respectively, the equations of motion of the particles are (cf. Fig. 6·1)

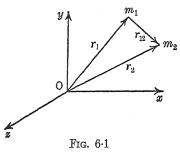
$$m_{1}\ddot{\mathbf{r}}_{1} = \frac{Gm_{1}m_{2}\mathbf{r}_{12}}{r_{12}^{3}},$$

$$m_{2}\ddot{\mathbf{r}}_{2} = \frac{Gm_{1}m_{2}\mathbf{r}_{21}}{r_{21}^{3}}.$$
(6.3-1)

Here G is the constant of gravitation (Sec. 2.3) and

$$\mathbf{r}_{12} = \mathbf{r}_2 - \mathbf{r}_1 \tag{6.3-2}$$

figure. Note that $r_{21} = -r_{12}$.



is the vector joining m_1 and m_2 , with direction as indicated in the The direction of the force on m_1 is along the line joining it to m_2 and directed toward m_2 , and similarly for m_2 . We can now apply the general theorems of Sec. 6.1, but in this case it is simpler to proceed somewhat differently since by doing so we can actually solve the problem completely. We decide to treat the problem as one of motion of m_1 relative to m_2 . Thus sub-

tract the first equation in (6.3-1) from the second after dividing by m_1 and m_2 , respectively. The result is

$$\ddot{\mathbf{r}}_{12} = \frac{-K}{r_{12}^3} \left(\frac{1}{m_1} + \frac{1}{m_2} \right) \mathbf{r}_{12}, \tag{6.3-3}$$

or

$$\mu \ddot{\mathbf{r}}_{12} = \frac{-K\mathbf{r}_{12}}{r_{12}^3},$$
(6.3–3')

where

$$\mu = \frac{m_1 m_2}{m_1 + m_2}, \quad K = G m_1 m_2.$$
 (6.3-4)

We see that eq. (6.3-3') describes the motion of a single particle of mass μ in an inverse square central force field having the same force constant K as that prevailing in the interaction of the two particles. Effectively, we have reduced the two-particle problem to a one-particle problem, and all the equations developed in Sec. 3.7 for inverse square central field motion about a fixed center will apply to the present case if we merely replace the mass wherever it occurs by the quantity μ and replace the constant c (eq. 3.7-1) by K. It is customary to call μ the reduced mass for the system of two particles. Clearly if $m_2 \gg m_1$, as would be the case if m_2 represented the mass of the sum and m_1 that of a planet in the solar system, $\mu \doteq m_1$.

It is often advantageous to refer the motion of the two particles m_1 and m_2 to neither mass but rather to the center of mass of the

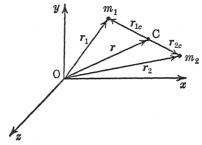


Fig. 6.2

system. In Fig. 6-2 we denote the center of mass by C with position vector $\bar{\mathbf{r}}$. The position vector of m_1 relative to C as origin is

$$\mathbf{r}_{1c} = \mathbf{r}_1 - \bar{\mathbf{r}},$$
 (6.3–5)

and correspondingly that of m_2 is

$$\mathbf{r}_{2c} = \mathbf{r}_2 - \bar{\mathbf{r}}. \tag{6.3-6}$$

Hence we have

$$\ddot{\mathbf{r}}_{1} = \ddot{\mathbf{r}} + \ddot{\mathbf{r}}_{1c} = \ddot{\mathbf{r}}_{1c},$$

$$\ddot{\mathbf{r}}_{2} = \ddot{\mathbf{r}} + \ddot{\mathbf{r}}_{2c} = \ddot{\mathbf{r}}_{2c},$$
(6.3-7)

since $\ddot{r} = 0$, there being no force present save the gravitational interaction. Moreover by the use of $(6\cdot3-5)$ and $(6\cdot3-6)$ and the definition of center of mass,

$$\mathbf{r}_{1c} = \frac{-m_2 \mathbf{r}_{12}}{m_1 + m_2}, \quad \mathbf{r}_{2c} = \frac{-m_1 \mathbf{r}_{21}}{m_1 + m_2}.$$
 (6.3-8)

Substitution into (6·3-3') gives the equations of motion in terms of \mathbf{r}_{1c} and \mathbf{r}_{2c}

$$\mu\ddot{\mathbf{r}}_{1c} = \frac{-m_2{}^3K\mathbf{r}_{1c}}{(m_1 + m_2){}^3r_{1c}{}^3},$$

$$\mu\ddot{\mathbf{r}}_{2c} = \frac{-m_1{}^3K\mathbf{r}_{2c}}{(m_1 + m_2){}^3r_{2c}{}^3}.$$
(6.3-9)

These may obviously be presented in the alternative form

$$m_{1}\ddot{\mathbf{r}}_{1c} = \frac{-m_{2}{}^{\circ}K\mathbf{r}_{1c}}{(m_{1} + m_{2})^{2}r_{1c}{}^{3}},$$

$$m_{2}\ddot{\mathbf{r}}_{2c} = \frac{-m_{1}{}^{2}K\mathbf{r}_{2c}}{(m_{1} + m_{2})^{2}r_{2c}{}^{3}}.$$

$$(6.3-10)$$

These differ from the original equations (6·3–1) only in the fact that the force constant is different for the two equations. We conclude that the path of each particle relative to the center of mass is of the same general geometrical form as that relative to the other particle. The center of mass itself, of course, either remains at rest or moves with constant speed in a straight line.

If $m_2 \gg m_1$, we have seen that $\mu \doteq m_1$ very nearly, while $r_{2c} \to 0$ and $r_{1c} \to r_{12}$. This means that when one mass is very much greater than the other, the motion of the heavier mass can be neglected compared with that of the other, and the problem reduces approximately to that of inverse square central field motion about a fixed center. The center of mass of the system in this case is very close to the heavier mass. Thus in the case of the motion of the two-particle system of the earth and the sun, the center of mass lies within the sun's surface and the approximation which considers merely the revolution of the earth about an effectively stationary sun is for many practical purposes a very good one.

In Sec. 3.9 we discussed the Bohr atomic model of hydrogen and assumed that the electron moves about a stationary nucleus. Since the mass of the nucleus is about 1840 times that of the electron, this is a rather good approximation. Nevertheless, the formula (3.9–14) for the semi-major axis of the orbits permitted by the quantum conditions can readily be corrected to take account of the actual motion of the nucleus. From what we have just learned in the present section it is only necessary to replace the mass of the electron by the reduced mass μ where now

$$\frac{1}{\mu}=\frac{1}{m_e}+\frac{1}{m_n},$$

in which m_e is the mass of the electron, and m_n that of the nucleus.

The formula (3.9-14) then becomes

$$a = \frac{(n_1 + n_2)^2 h^2}{4\pi^2 \left(\frac{m_e m_n}{m_c + m_n}\right) e^2}$$
 (6·3–11)

Since

$$\mu \sim m_e \left(1 - \frac{1}{1840}\right)$$
,

the correction is a small one. Nevertheless the modification has been detected experimentally by precise spectroscopic measurements.¹

6.4. Torque and Moment of Momentum. Let us now revert to the equation of motion of a system of n particles, i.e. (6.1-2), which we repeat here for convenience

$$m_j \ddot{\mathbf{r}}_j = \mathbf{F}_{je} + \sum_{k(\neq j)=1}^n \mathbf{F}_{jk}.$$

We now multiply both sides of the equation by the position vector \mathbf{r}_{j} , but we shall not do it by the *dot* product defined earlier. Rather we define a new kind of vector called the *cross* product.

If we have two vectors A and B, we shall define the vector or cross product as follows

$$\mathbf{A} \times \mathbf{B} = \epsilon A B \sin \theta, \qquad (6.4-1)$$

where A and B are as usual the magnitudes of the vectors A and B respectively. The angle θ is the smaller of the two angles between the positive directions of A and B, and ϵ is a unit vector (i.e. $|\epsilon| = 1$) perpendicular to the plane of A and B and directed in such a way that as one looks along it a clockwise rotation will carry A into B. The cross product is therefore a vector whose magnitude is the area of the parallelogram of which A and B are two adjacent sides. We note that the vector product does not satisfy the commutative law. Rather

$$A \times B = -B \times A. \tag{6.4-2}$$

However it can be shown that the distributive law holds, i.e.,

$$A \times (B + C) = A \times B + A \times C. \tag{6.4-3}$$

¹Cf. Ruark and Urey, "Atoms, Molecules and Quanta," p. 126.

If we apply the definition to the unit vectors i, j, k we get the interesting and often useful results

$$\mathbf{i} \times \mathbf{i} = \mathbf{j} \times \mathbf{j} = \mathbf{k} \times \mathbf{k} = 0,$$

 $\mathbf{i} \times \mathbf{j} = \mathbf{k}, \quad \mathbf{j} \times \mathbf{k} = \mathbf{i}, \quad \mathbf{k} \times \mathbf{i} = \mathbf{j}.$ (6.4-4)

Moreover

$$\frac{d}{dt}(\mathbf{A} \times \mathbf{B}) = \dot{\mathbf{A}} \times \mathbf{B} + \mathbf{A} \times \dot{\mathbf{B}}, \qquad (6.4-5)$$

which can be demonstrated by expanding A and B in terms of their components along the x, y, z axes and using $(6 \cdot 4-4)$.

We are now ready to apply the cross product to the present problem by forming

$$\mathbf{r}_j \times \ddot{\mathbf{r}}_j = \epsilon_j r_j |\ddot{\mathbf{r}}_j| \sin(\mathbf{r}_j, \ddot{\mathbf{r}}_j).$$
 (6.4-6)

Here $(\mathbf{r}_j, \ddot{\mathbf{r}}_j)$ is the smaller of the angles between the positive directions of the vectors \mathbf{r}_j and $\ddot{\mathbf{r}}_j$ (cf. θ in Fig. 6.3), and $\boldsymbol{\epsilon}_j$ is the unit vector normal to the plane of \mathbf{r}_j and $\ddot{\mathbf{r}}_j$ and in such a direction that, as one looks along it, a clockwise rotation will carry \mathbf{r}_j into

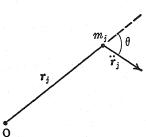


Fig. 6.3

 $\ddot{\mathbf{r}}_j$. The quantity $\mathbf{r}_j \times \ddot{\mathbf{r}}_j$ will be defined as the *moment* of the acceleration vector about the axis through O perpendicular to the plane of \mathbf{r}_j and $\ddot{\mathbf{r}}_j$. Similarly

$$\mathbf{r}_{j} \times \mathbf{F}_{je}$$

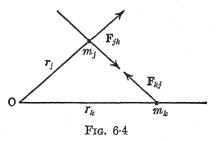
will be defined as the moment of the force F_{je} about the axis through the origin perpendicular to the plane of r_j and F_{je} . The moment of a force about

an axis is also termed the *torque* due to the force about the axis in question. We shall often refer to it more simply as the torque about the point through which the axis passes. It is, of course, a vector and is usually designated by the symbol L. The result of taking the moments of both sides of the equation of motion then appears as

$$m_j \mathbf{r}_j \times \ddot{\mathbf{r}}_j = \mathbf{L}_{je} + \sum_{k(\neq j)=1}^n \mathbf{L}_{jk},$$
 (6.4-7)

where L_{je} is the torque due to the external force on the mass m_{j} ,

while L_{jk} is the torque due to the interaction force between m_j and m_k . Let us assume that this interaction force is directed along the straight line joining m_j and m_k . It follows from Fig. 6.4 that the torque due to the force F_{jk} on m_j about the normal axis through



O is equal and opposite in direction to that due to the force F_{kj} on m_k , about the same axis. Hence if we sum eq. (6.4-7) over all the particles of the system the double sum

$$\sum_{j=1}^{n} \sum_{k(\neq j)=1}^{n} \mathbf{L}_{jk}$$

will vanish, and we shall have simply

$$\sum_{j=1}^{n} m_{j} \mathbf{r}_{j} \times \ddot{\mathbf{r}}_{j} = \sum_{j=1}^{n} \mathbf{L}_{je}.$$
 (6.4-8)

To interpret this more conveniently, we use eq. (6.4-5) and finally get

$$\frac{d}{dt}[\mathbf{r}_j \times (\mathbf{m}_j \dot{\mathbf{r}}_j)] = \mathbf{r}_j \times (m_j \ddot{\mathbf{r}}_j), \qquad (6.4-9)$$

whence (6.4-8) becomes

$$\frac{d}{dt}\sum_{j=1}^{n}\mathbf{r}_{j}\times(m_{j}\dot{\mathbf{r}}_{j})=\sum_{j=1}^{n}\mathbf{L}_{je}.$$
 (6.4-10)

The quantity $\mathbf{r}_j \times (m_j \dot{\mathbf{r}}_j)$ from its method of construction is naturally termed the moment of momentum of the jth particle about the origin. Eq. (6·4–10) then says that the time rate of change of the total moment of momentum of the collection of particles about the origin is equal to the total torque or force moment produced by the external forces about the same point. The reader will note its analogy with the force equation (6·1–9). Corresponding to the theorem of the conservation of momentum in Sec. 6·1 we now

have the theorem of the conservation of moment of momentum: if the total external torque on a system of particles about a point vanishes, the total moment of momentum of the system about this point is constant. It must be remembered, of course, that the validity of this result depends on the internal interaction forces being directed along the lines joining the particles respectively. Otherwise the torques due to the internal forces will not cancel out. An illustration of the theorem is found in the solar system in so far as the whole system is not acted on by external forces. Another illustration is found in the electron motions in the planetary atom model already described in Chapter III. Here, of course, it is only approximate since internal magnetic forces produce torques which do not entirely cancel out. Moreover, the electrons themselves are now assumed to have spins which must be taken into account.

The reader should show for his own satisfaction that eq. (6·4–10) when written in terms of rectangular coördinates becomes

$$\frac{d}{dt} \sum_{j=1}^{n} m_{j} (y_{j} \dot{z}_{j} - z_{j} \dot{y}_{j}) = \sum_{j=1}^{n} (y_{j} F_{jez} - z_{j} F_{jey}),$$

$$\frac{d}{dt} \sum_{j=1}^{n} m_{j} (z_{j} \dot{x}_{j} - x_{j} \dot{z}_{j}) = \sum_{j=1}^{n} (z_{j} F_{jez} - x_{j} F_{jez}), \quad (6.4-11)$$

$$\frac{d}{dt} \sum_{j=1}^{n} m_{j} (x_{j} \dot{y}_{j} - y_{j} \dot{x}_{j}) = \sum_{j=1}^{n} (x_{j} F_{jey} - y_{j} F_{jez}).$$

Can we bring the center of mass into this picture? In line with the discussion of the motion of two particles let us refer the motion of the whole collection of n particles to the center of mass. Thus we write the position vector of the jth particle in the form (analogous to the special case of eq. (6.3-5))

$$\mathbf{r}_{j} = \bar{\mathbf{r}} + \mathbf{r}_{jc}, \tag{6.4-12}$$

where $\bar{\mathbf{r}}$ is the position vector of the center of mass, and \mathbf{r}_{jc} is the position vector of m_j relative to the center of mass. The fundamental equation (6·1-9) becomes

$$\frac{d}{dt}\left(m\dot{\mathbf{r}}\right) = \sum \mathbf{F}_{je}, \qquad (6.4-13)$$

since by definition $\sum_{j=1}^{n} m_j r_{jc} = 0$. Eq. (6.4-13) is, of course,

equivalent to eq. $(6\cdot1-8)$. This is nothing new, then. However, let us introduce the substitution $(6\cdot4-12)$ into $(6\cdot4-10)$. This gives

$$\frac{d}{dt}\sum_{j=1}^{n}m_{j}(\mathbf{r}+\mathbf{r}_{jc})\times(\dot{\mathbf{r}}+\dot{\mathbf{r}}_{jc})=\sum_{j=1}^{n}(\mathbf{r}+\mathbf{r}_{jc})\times\mathbf{F}_{jc}.$$

After expansion and use of the definition of the center of mass and (6·4–13) the result is

$$\frac{d}{dt} \sum_{j=1}^{n} \mathbf{r}_{jc} \times (m_j \dot{\mathbf{r}}_{jc}) = \sum_{j=1}^{n} \mathbf{r}_{jc} \times \mathbf{F}_{je}. \tag{6.4-14}$$

This says that the time rate of change of the total moment of momentum about the center of mass is equal to the total torque about the center of mass. In other words we can completely describe the motion of the collection of particles by the motion of the center of mass (6·4–13) and the motion relative to the center of mass (6·4–14). This plan turns out to be particularly appropriate when the collection of particles forms a rigid body (Chap. VII). Another important application is met with in nuclear physics where in the theory of the deuteron (the nucleus of the hydrogen isotope of mass two), for example, it is convenient to separate the motion of the center of mass of the two particles (i.e., the neutron and the proton) from their relative motion.

6.5. Energy of a System of Particles. Let us revert once more to the equation of motion of the jth particle and form the dot product of both sides by $\dot{\mathbf{r}}_j dt$ to find the work done by the forces in time dt. Thus

$$m_j \dot{\mathbf{r}}_j \cdot \ddot{\mathbf{r}}_j dt = \mathbf{F}_{je} \cdot \dot{\mathbf{r}}_j dt + \sum_{k(\neq j)=1}^n \mathbf{F}_{jk} \cdot \dot{\mathbf{r}}_j dt.$$
 (6.5-1)

The left-hand side becomes

$$\frac{1}{2}m_jd(v_j{}^2),$$

where $v_j^2 = \dot{\mathbf{r}}_j \cdot \dot{\mathbf{r}}_j$ is the square of the resultant velocity of the *j*th particle. Moreover, since we can write $\dot{\mathbf{r}}_j dt = d\mathbf{r}_j$, we have finally on summing over j

$$d\sum_{j=1}^{n} \left(\frac{1}{2}m_{j}v_{j}^{2}\right) = \sum_{j=1}^{n} \mathbf{F}_{je} \cdot d\mathbf{r}_{j} + \sum_{j=1}^{n} \sum_{k(\neq j)=1}^{n} \mathbf{F}_{jk} \cdot d\mathbf{r}_{j}.$$
 (6.5–2)

Integrating over the path from r_j' to r_j" gives

$$K_{1} - K_{0} = \int_{\mathbf{r}_{j'}}^{\mathbf{r}_{j''}} \left[\sum_{i} \mathbf{F}_{je} \cdot d\mathbf{r}_{j} + \sum_{j=1}^{n} \sum_{k(\neq j)=1}^{n} \mathbf{F}_{jk} \cdot d\mathbf{r}_{j} \right]$$
(6.5-3)

This is the generalization of the work-energy theorem to the case of a collection of particles: the total work done by the external and internal forces during the motion of the system is equal to the change in the total kinetic energy. The double sum on the right can be written in much simpler form if we assume that the force \mathbf{F}_{jk} is directed along the line joining the jth and kth particles and is a function only of the distance between them and moreover the same function for all particles. Thus let us suppose that

$$\mathbf{F}_{jk} = f(r_{jk})\mathbf{r}_{jk}, \tag{6.5-4}$$

where

$$\mathbf{r}_{jk} = \mathbf{r}_j - \mathbf{r}_k.$$

Now the double sum consists of pairs of terms of the form

$$f(r_{jk})\mathbf{r}_{jk} \cdot d\mathbf{r}_{j} + f(r_{kj})\mathbf{r}_{kj} \cdot d\mathbf{r}_{k}$$

$$= f(r_{jk})\mathbf{r}_{jk} \cdot d\mathbf{r}_{jk}. \tag{6.5-5}$$

Hence (6.5-3) becomes

$$K_{1} - K_{0} = \int_{\mathbf{r}_{j}'}^{\mathbf{r}_{j}''} \left[\sum_{j=1}^{n} \mathbf{F}_{je} \cdot d\mathbf{r}_{j} + \sum f(r_{jk}) \mathbf{r}_{jk} \cdot d\mathbf{r}_{jk} \right], \quad (6.5-6)$$

where in the second sum we include all pairs of j and k with j < kand k running from 1 to n. Now suppose there exists a potential energy function V_{je} such that

$$\sum_{i} \mathbf{F}_{je} \cdot d\mathbf{r}_{j} = -\sum_{i} \left(\frac{\partial V_{je}}{\partial x_{j}} dx_{i} + \frac{\partial V_{je}}{\partial y_{j}} dy_{j} + \frac{\partial V_{je}}{\partial z_{j}} dz_{j} \right) = -\sum_{j} dV_{je}.$$
Moreover, accuracy that

Moreover, assume that

$$\sum f(r_{jk})\mathbf{r}_{jk} \cdot d\mathbf{r}_{jk} = -\sum \left[\frac{\partial V(r_{jk})}{\partial x_{jk}} dx_{jk} + \frac{\partial V(r_{jk})}{\partial y_{jk}} dy_{jk} + \frac{\partial V(r_{jk})}{\partial z_{jk}} dz_{jk} \right]$$
$$= -\sum dV(r_{jk}), \tag{6.5-8}$$

where $x_{jk} = x_j - x_k$, etc. Eq. (6.5-6) then becomes

$$K_1 + \sum_j V_{je}^{\prime\prime} + \sum_{k,j} V^{\prime\prime}(r_{jk}) = K_0 + \sum_j V_{je}^{\prime} + \sum_{k,j} V^{\prime}(r_{jk}).$$
 (6.5–9)

We are then led to call $\sum_{j} V_{je}$ the total potential energy of the external forces, and $\sum V(r_{jk})$ the total potential energy of the internal forces. The total energy of the whole system remains constant: this is the theorem of the conservation of mechanical energy for a system of particles. From (6.5-8) it should be clear that if the mutual interaction forces are central forces, i.e., act along the lines joining the particles and vary only with the interparticle distance, the potential energy function $V(r_{jk})$ always exists. Hence in a system of particles acted on only by mutual forces of the central type, the total energy remains constant.

An important fact to note from (6.5–9) is that the internal forces make a contribution to the potential energy and hence appear in the energy conservation equation, whereas they cancel out in the momentum and moment of momentum conservation equations. This is well brought out in the energy equation for the two-particle problem discussed in Sec. 6.3. Let us consider this special case.

If we dot multiply the first equation in (6.3-1) by $\dot{\mathbf{r}}_1 dt = d\mathbf{r}_1$ and the second by $\dot{\mathbf{r}}_2 dt = d\mathbf{r}_2$, and add, by the arguments used in going from (6.5-1) to (6.5-2), we get

$$d\left(\frac{1}{2}m_{1}v_{1}^{2} + \frac{1}{2}m_{2}v_{2}^{2}\right) = \frac{Gm_{1}m_{2}}{r_{12}^{3}}\mathbf{r}_{12} \cdot (d\mathbf{r}_{1} - d\mathbf{r}_{2})$$

$$= -\frac{Gm_{1}m_{2}}{r_{12}^{3}}\mathbf{r}_{12} \cdot d\mathbf{r}_{12}, \qquad (6.5-10)$$

where the last step comes from (6·3–2). Integration of both sides leads to

$$\frac{1}{2}m_1v_1^2 + \frac{1}{2}m_2v_2^2 = -Gm_1m_2\int \frac{\mathbf{r}_{12}}{r_{12}^3} \cdot d\mathbf{r}_{12}$$

$$= -Gm_1m_2\int \frac{dr_{12}}{r_{12}^2}, \qquad (6.5-11)$$

since

$$\mathbf{r}_{12} \cdot d\mathbf{r}_{12} = r_{12} dr_{12}.$$
 (6.5–12)

We therefore finally obtain as the energy equation

$$\frac{1}{2}m_1v_1^2 + \frac{1}{2}m_2v_2^2 - \frac{Gm_1m_2}{r_{12}} = U, (6.5-13)$$

where U is the total energy. The potential energy of interaction of the two particles is

$$V_{12} = -\frac{Gm_1m_2}{r_{12}}, (6.5-14)$$

as might have been expected from the discussion leading up to eq. $(4\cdot4-2)$. Here there is no external force and hence no external potential energy. In $(6\cdot5-13)$ the speeds v_1 and v_2 refer, of course, to the fixed origin. We can, if we please, express the energy equation in terms of the relative velocity $\dot{\mathbf{r}}_{12}$. In fact this is a simpler procedure. From eq. $(6\cdot3-3')$ we have at once

$$\frac{\mu v_{12}^2}{2} - \frac{Gm_1m_2}{r_{12}} = U, (6.5-15)$$

in which now the reduced mass once more appears. We can also express the energy equation in terms of the motion relative to the center of mass. By the use of $(6\cdot3-5)$ and $(6\cdot3-6)$ we can write for the left-hand side of $(6\cdot5-13)$

$$\frac{1}{2}m_1v_1^2 + \frac{1}{2}m_2v_2^2 = \frac{1}{2}(m_1 + m_2)\dot{\mathbf{r}} \cdot \dot{\mathbf{r}} + \frac{1}{2}m_1\dot{\mathbf{r}}_{1c} \cdot \dot{\mathbf{r}}_{1c}
+ \frac{1}{2}m_2\dot{\mathbf{r}}_{2c} \cdot \dot{\mathbf{r}}_{2c}.$$
(6.5-16)

This says that the total kinetic energy in the fixed system of reference is equal to the kinetic energy of the center of mass (with the whole mass $m_1 + m_2$ concentrated there) plus the kinetic energy of the two particles relative to the center of mass. This scheme is often a useful one and of course may be generalized to apply to a collection of an arbitrary number of particles. We shall make good use of it in the study of rigid bodies in Chap. VII.

In the case of the collisions of spherical particles discussed in Sec. 6.2 it is extremely difficult if not practically impossible to express the interaction forces in terms of readily measurable quantities. Hence we cannot readily find an expression for the potential energy at every instant during the impact and therefore cannot write the energy equation for the interaction. However, if we are content to use the Newtonian coefficient of restitution,

it is possible to examine the change in kinetic energy brought about by the collision.

The total kinetic energy of the system before collision is

$$K_u = \frac{1}{2}m_1u_1^2 + \frac{1}{2}m_2u_2^2, \tag{6.5-17}$$

whereas that after collision is

$$K_v = \frac{1}{2}m_1v_1^2 + \frac{1}{2}m_2v_2^2. \tag{6.5-18}$$

The change in kinetic energy is

$$\Delta K = K_v - K_u. \tag{6.5-19}$$

We can evaluate this by direct substitution for v_1 and v_2 from $(6\cdot2-4)$, so as to express ΔK in terms of u_1 and u_2 only. However, this turns out to be algebraically tedious, and we find it simpler to avail ourselves of the result demonstrated just above, namely, that the total kinetic energy is equal to that of the center of mass plus the kinetic energy relative to the center of mass. Since the velocity of the center of mass cannot be changed by the collision, the whole change in the kinetic energy of the system is the change in the kinetic energy relative to the center of mass. Now, using the notation of Sec. 6·3, we have

$$m_1u_{1c} + m_2u_{2c} = m_1v_{1c} + m_2v_{2c} = 0,$$
 (6.5-20)

whence from (6.2-3) we conclude that

$$v_{1c} = -eu_{1c}, \quad v_{2c} = -eu_{2c}.$$
 (6.5–21)

Therefore the change in kinetic energy becomes simply

$$\Delta K = \frac{1}{2}m_1v_{1c}^2 + \frac{1}{2}m_2v_{2c}^2 - \frac{1}{2}m_1u_{1c}^2 - \frac{1}{2}m_2u_{2c}^2$$

$$= \frac{1}{2}m_1(e^2 - 1)u_{1c}^2 + \frac{1}{2}m_2(e^2 - 1)u_{2c}^2.$$
 (6.5–22)

From the fact that the coefficient of restitution e is a proper fraction (≤ 1) it follows that

$$\Delta K \le 0. \tag{6.5-23}$$

Unless e=1, the total kinetic energy of the system of spherical particles decreases on collision even though the total momentum remains constant. The kinetic energy thus lost reappears in the form of heat and sound, as is clearly shown by experiment.

6.6. The Virial of Clausius. Closely related to the energy of a system of particles is the *virial* of Clausius. Consider the following

158

scalar function of the position vectors of the n particles of the system

$$M = \sum_{j=1}^{n} \frac{1}{2} m_{j} \mathbf{r}_{j} \cdot \mathbf{r}_{j}. \tag{6.6-1}$$

Differentiating this twice with respect to the time gives

$$\ddot{M} = \sum_{j=1}^{n} m_j \dot{\mathbf{r}}_j \cdot \dot{\mathbf{r}}_j + \sum_{j=1}^{n} m_j \mathbf{r}_j \cdot \ddot{\mathbf{r}}_j. \tag{6.6-2}$$

Now \dot{M} is a function of the time. If we integrate it with respect to the time between the limits 0 and τ , where τ is an arbitrarily long time, and divide by τ , we shall obtain the *time average* of M over the time interval in question. Thus, if we use a bar to denote average with respect to time, we have

$$\overline{\ddot{M}} = \frac{1}{\tau} \int_0^{\tau} \dot{M} dt = \frac{1}{\tau} (\dot{M}_{\tau} - \dot{M}_0),$$
 (6.6–3)

where, of course, from (6.6-1)

$$\dot{M} = \sum_{j=1}^{n} m_j \mathbf{r}_j \cdot \dot{\mathbf{r}}_j, \qquad (6.6-4)$$

and \dot{M}_{τ} and \dot{M}_{0} refer to the values of \dot{M} at $t=\tau$ and t=0, respectively. If the particles are assumed to move always in a closed space, the magnitude of \mathbf{r}_{j} can never exceed a certain maximum value. We shall assume, moreover, that the velocities $\dot{\mathbf{r}}_{j}$ are likewise bounded, which is quite reasonable for a collection of particles for which the kinetic energy cannot become infinitely great. Hence \dot{M} is bounded and as τ increases

$$\overline{\dot{M}} \rightarrow 0.$$
 (6.6–5)

Therefore from (6.6-2)

$$\frac{1}{2} \sum_{j=1}^{n} m_j \dot{\mathbf{r}}_j \cdot \dot{\mathbf{r}}_j = -\frac{1}{2} \sum_{j=1}^{n} m_j \mathbf{r}_j \cdot \ddot{\mathbf{r}}_j.$$
 (6.6–6)

Now the left-hand side of (6.6-6) is the time average of the total kinetic energy of the system. The right-hand side may be rewritten by recalling that $m_j \ddot{\mathbf{r}}_j$ is the resultant force \mathbf{F}_j acting on the jth particle. We may then write

$$\Omega = -\frac{1}{2} \sum_{j=1}^{n} \mathbf{r}_{j} \cdot \mathbf{F}_{j}. \tag{6.6-7}$$

The quantity Ω was called by Clausius the *virial* of the system. Eq. (6·6-6) expresses the result that the average kinetic energy of the system of particles moving subject to the imposed conditions is equal to its virial. This is the virial theorem.

An interesting application of the virial theorem is the relation between the average kinetic energy and average potential energy of a system in which the only forces acting are central forces between the various particles. We shall illustrate this for the simple case of a single particle subject to a central attractive force varying as the nth power of the distance to a force center. Then $(6\cdot6-6)$ reduces (with $\mathbf{F} = Cr^{n-1}\mathbf{r}$) to

$$\overline{K} = -\frac{1}{2} \overline{\mathbf{F} \cdot \mathbf{r}} = -\frac{1}{2} C \overline{r^{n-1}}.$$

But from (4.3-8) the potential energy corresponding to this type of central field is

$$V(r) = -\frac{Cr^{n+1}}{n+1},$$

and consequently

$$\bar{K} = \frac{n+1}{2} \cdot \bar{V}. \tag{6.6-8}$$

In particular, for the inverse square field, where n = -2, the result becomes

$$\bar{K} = -\frac{\bar{V}}{2} \cdot \tag{6.6-9}$$

From (6.6-9) the total energy becomes

$$E = \frac{\overline{V}}{2}, \qquad (6.6-10)$$

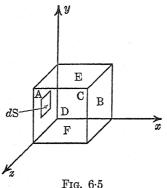
which is, of course, negative, corresponding to the fact that the force is here a binding force. The generalization to a collection of particles is left as an exercise.

6.7. Elementary Kinetic Theory of Gases. Equation of State of an Ideal Gas. One of the most important examples of a system of particles is provided by the kinetic theory, which considers a gas to be composed of a very large number of very small particles called molecules moving with varying velocities in every direction, colliding with each other and rebounding with undiminished speed

from the walls of the vessel in which the gas is confined. In the so-called ideal gas the molecules are assumed to exert no forces on each other and consequently the only forces on them are due to the impact on the walls. Let us calculate the virial for an ideal gas consisting of N molecules confined in a cubical box of side a. We suppose the gas to be in a state of equilibrium, i.e., assume that the spatial distribution of the molecules within the box is uniform so that the density is the same in all parts of the vessel and does not change with the time. It will be convenient to expand the scalar product in (6.6-7) and express the virial in terms of rectangular components. Thus

$$\Omega = -\frac{1}{2} \sum_{j=1}^{N} (x_j F_{jx} + y_j F_{jy} + z_j F_{jz}).$$
 (6.7-1)

Now the forces \mathbf{F}_{iz} , etc., have non-vanishing values only for the values of x_i , y_i , z_i at the walls. These forces arise from the change in momentum experienced by the molecules in their reflection from the walls. It might seem, at first thought, difficult to compute them. However, when we recall that it is the time average



which is involved in (6.7-1), we reflect that the average wall forces may be replaced by the integrated pressure p (force per unit area) which on the postulates of the kinetic theory is assumed to be due to the average effect of the continual bombardment of the walls by the molecules. We agree to establish coördinate axes as in Fig. 6.5. To form $\sum x_i F_{ix}$ we must multiply the x coördinate of each molecule by the x component of the

force acting on it, sum over all molecules, and then average. Now the only contribution to $\sum x_i F_{ix}$ will come from the walls A and B. That at wall A must vanish since x is zero at all points of this wall. At B we may assume that each single molecule is acted on by a surface element dS. Hence $\sum x_i F_{ix}$ becomes

$$-pa \int dS = -pa^3, \qquad (6.7-2)$$

for at B, $x_i = a$ everywhere and the integration must be taken over the whole wall. The negative sign results from the fact that the wall force acts inward. The contribution of the pair of walls A and B to $\sum y_i F_{iy}$ and $\sum z_i F_{iz}$ will be zero since these walls are normal to the x axis and the force they exert is directed wholly along the x axis. Proceeding similarly we find that the contribution to $\sum y_i F_{iy}$ of the pair of walls E, F is $-pa^3$ and the same is true of the contribution of C, D to $\sum z_i F_{iz}$. Hence the total virial due to the forces exerted by the walls becomes

$$\Omega = \frac{3}{2}pV_0, \qquad (6.7-3)$$

where $V_0 = a^3$ is the volume of the cubical box. The virial theorem (6.6-6) then gives

$$\frac{1}{2} \sum m_j v_j^2 = \frac{3}{2} p V_0, \tag{6.7-4}$$

where v_j is the resultant velocity of the jth molecule.

The formula (6.7-4) at once gains fundamental importance for the thermal properties of gases if we assume that the temperature of an ideal gas is characterized entirely by the total average kinetic energy. Then constant temperature implies constant average kinetic energy. On this assumption (6.7-4) says that the product of pressure and volume for an ideal gas at constant temperature is constant. This is Boyle's law.

It may be remarked that the foregoing derivation can be carried through much more generally using a vessel of any size or shape. The reader may here refer to Lindsay, "Physical Statistics" (John Wiley, New York, 1941) p. 70, for the more general treatment. Also consult Loeb, "The Kinetic Theory of Gases." (McGraw-Hill, N.Y., 1927.)

Let us examine eq. (6.7-4) more closely. Suppose that the molecules are all of equal mass and moreover suppose that v_i^2 is the same for all. This is equivalent (for this special case) to the assumption that the mean kinetic energy is the same for all the molecules, i.e., that there is equipartition of energy among them. This hypothesis or one closely connected with it has played a significant role in the development of the kinetic theory.1

¹ See again, Lindsay, op. cit., p. 67.

Eq. (6.7-4) now becomes (with $\overline{v_i^2} = \overline{v^2}$)

$$pV_0 = \frac{1}{3}mN\overline{v^2}, (6.7-5)$$

where N is the total number of molecules in the volume V_0 , and m is the mass of each. It should be carefully noted that $\overline{v^2}$ represents the time average of the square of the velocity, i.e., it has no connection with direction. If we were to consider, for example, the average velocity, then we should have to admit that it must be zero, since for every molecule traveling at a given instant in a definite direction with a certain speed there will be to a very high degree of probability another one moving in the opposite direction with the same speed. Now if we divide through eq. (6.7-5) by V_0 , we have

$$p = \frac{1}{3}mn\overline{v^2},\tag{6.7-6}$$

where n is the number of molecules per unit of volume. Hence $mn = \rho$, where ρ is the density of the gas, and we write finally

$$p = \frac{1}{3}\rho \overline{v^2}.\tag{6.7-7}$$

This provides at once a means of calculating the mean square velocity of the molecules and its square root, called the *root mean* square velocity. Thus

$$\sqrt{\overline{v^2}} = v_m = \sqrt{\frac{3p}{\rho}} \cdot \tag{6.7-8}$$

For oxygen at standard pressure (76 cm. of mercury) and at temperature 0° C., $\rho=0.001429~{\rm gram/cm^3}$, and hence we calculate

$$v_m = 461.2 \text{ meters/sec},$$

approximately. The reader should calculate for himself the corresponding values for other gases. For example, he will find it of interest to work out the value for hydrogen and thus obtain a clue to the reason why there is so little hydrogen in the earth's atmosphere (cf. Sec. 2·3). The foregoing is sufficient to indicate the order of magnitude of molecular velocities as calculated from kinetic theory.

Let us now introduce a more exact connection between the kinetic energy of the molecules and the temperature, and assume

$$mv_m^2 = 3kT, (6.7-9)$$

where T is the absolute temperature. It then develops from a study of the experimental data on many gases that k is a *universal* constant, depending only on the temperature scale used. It is usually denoted as the Boltzmann gas constant or gas constant per molecule. If we knew the mass of the oxygen molecule, for example, we could calculate the value of k from (6.7-9). The most reliable data yield $k = (1.3708 \pm 0.0014) \times 10^{-16}$ ergs/degree C. Now substituting into (6.7-6) we have

$$p = nkT, (6.7-10)$$

or multiplying through by a volume V_0 ,

$$pV_0 = NkT. (6.7-11)$$

Eq. (6·7–11) is the general gas equation for an ideal gas, i.e., one in which the mutual forces of the molecules may be neglected compared with the forces of the walls on the molecules. It is also frequently called the equation of state.

6.8. Some Consequences of the Kinetic Theory. Specific Heats of Gases. Several other results important for the theory of gases follow from the use of the virial theorem. We want to stress these here because they illustrate the very significant rôle that mechanical methods have played in the development of the theory of heat and the constitution of bodies.

In the first place note from eq. (6.7-11) that, since k is a universal constant, the number of molecules in volume V_0 of any ideal gas, given by

$$N = \frac{pV_0}{kT}, \qquad (6.8-1)$$

is dependent only on the pressure, volume and temperature of the gas. It follows at once that if we have two gases at the same temperature and pressure, equal volumes of the two will contain the same number of molecules. This is the celebrated hypothesis of Avogadro, so important for chemistry. It here follows directly from kinetic theory considerations. Incidentally we note that if we were to assume Avogadro's hypothesis in a priori fashion and combine it with (6.7-6) the result would be that $\frac{1}{3}mv^2$ must be the same for all gases at the same temperature. This at once suggests a kinetic theory definition of temperature like (6.7-9).

The independent evidence in favor of Avogadro's hypothesis thus lends the greater weight to an assumption of the form (6.7-9). Of course, we cannot go here into the detail that a close inspection of this subject demands. Nevertheless the reader should see in it an illustration of the progressive interplay of hypotheses in constructing physical theories for various types of physical phenomena. Another conclusion closely allied to that just arrived at follows at once. Since $nm = \rho$, we have for two different ideal gases

 $n_1 m_1 = \rho_1, \quad n_2 m_2 = \rho_2.$ (6.8-2)

But if the temperature and pressure are the same, $n_1 = n_2$, as we have just seen. Hence under these conditions

$$\frac{m_1}{m_2} = \frac{\rho_1}{\rho_2},\tag{6.8-3}$$

that is, for constant temperature and pressure the ratio of the densities of two gases is equal to the ratio of the masses of their constituent molecules (i.e., chemically speaking the ratio of their molecular weights). This law was discovered originally by the chemist Gay-Lussac.

If we can suppose that when a gas is allowed to effuse through a very small opening in a vessel the outward velocity is proportional to the root mean square velocity v_m , it follows from eqs. (6.7-9) and (6.8-3) that

$$\frac{v_{m1}}{v_{m2}} = \sqrt{\frac{m_2}{m_1}} = \sqrt{\frac{\rho_2}{\rho_1}} \tag{6.8-4}$$

for two gases at the same temperature and pressure. This, however, states in words that the effusion velocity (for gases at the same temperature and pressure) varies inversely as the square root of the gas density. This is actually satisfied experimentally to a considerable degree of approximation.

Some interesting results follow from the consideration of the energy of the molecules. In an ideal gas the energy of the gas is the sum of the separate energies of motion of its molecules. The latter energy for each molecule can be divided into two parts, viz., the kinetic energy of the molecule as a whole (i.e. the kinetic energy of the center of mass) and the kinetic energy of the separate parts of the molecule (supposing it has a structure) relative to the

center of mass. We shall assume that the latter energy, which we may call the internal kinetic energy of the molecule, is a constant fraction of the kinetic energy of translation. So we shall write for the total energy of a single molecule

$$E = \frac{1}{2}\beta m\overline{v^2},\tag{6.8-5}$$

where β is a constant which should be the same for gases of the same constitution. The energy for unit volume is then

$$E_V = \frac{1}{2}\beta nm\overline{v^2}, \qquad (6.8-6)$$

where n is the number of molecules per unit volume. Now from eq. (6.7-6) we may write this in the form

$$E_V = \frac{3}{2}\beta p,\tag{6.8-7}$$

p being the pressure of the gas. The energy per unit mass of the gas is then

$$E_m = \frac{3}{2} \beta \frac{p}{\rho}, \qquad (6.8-8)$$

where ρ is the density. Now the specific heat of a gas at constant volume c_V is measured by the increase in energy per unit mass per unit increase in absolute temperature while the volume remains the same. Hence we have

$$c_{V} = \left(\frac{dE_{m}}{dT}\right)_{V} = \frac{3}{2}\beta \left[\frac{d}{dT}\left(\frac{p}{\rho}\right)\right]_{V}.$$
 (6.8-9)

This is expressed in mechanical units, i.e., ergs per degree. To change to calories per degree one must divide by the mechanical equivalent of heat, $J=4.2\times 10^7$ ergs per calorie. What we are here interested in is, however, the fact that for an ideal gas (6.8-9) yields

$$c_V = \frac{3}{2} \frac{\beta k}{m} \cdot \tag{6.8-10}$$

This shows that for a given gas the specific heat at constant volume should be independent of temperature and pressure. This is experimentally true over a wide range, but as one might expect breaks down at low temperatures, i.e., under the conditions such that the interaction forces can no longer be neglected. From (6.8-10) it is clear that c_v is inversely proportional to the mass of a

molecule of the gas and directly proportional to β . If β were the same for a number of gases it would follow that the product of m by c_V or that which is proportional to this, the product of c_V by the molecular weight, is constant. This is the celebrated law of Dulong and Petit as applied to gases. It holds for a great many gases but, of course, is not valid for all gases under all conditions. It is also found to hold for most solid elements (the molecular weight being replaced by the atomic weight) if the temperature is high enough. The exact study of the specific heats of substances in general is very intricate. It has only yielded to successful treatment by the use of the quantum theory of the constitution of matter. 1

If the molecules are of non-vanishing size (and there are many lines of evidence to indicate that this is so) we must admit that with their very considerable velocities they will suffer frequent collisions. The time elapsing between successive collisions will naturally be quite varied. Nevertheless we can profitably imagine an average distance through which a molecule may travel before encountering another molecule. This distance is called the mean free path and is usually denoted by \(\lambda\). There are in kinetic theory several methods of estimating it. It would take us too far afield to enter into these here. However, it is reasonable to expect that the more molecules there are per unit volume, the shorter will be the mean free path, i.e., \(\lambda\) varies inversely as the density. should also expect the size of the molecules to affect the value of λ . again inversely. Calculation confirms these conjectures, and indeed Clausius obtained the following expression for the mean free path

$$\lambda = \frac{3}{4} \cdot \frac{1}{n\pi D^2}, \qquad (6.8-11)$$

where n is as before the number of molecules per unit volume (assumed approximately constant), and D is the diameter of a molecule considered as a rigid sphere. This is an idealized picture. Nevertheless it has proved extremely valuable.

6.9. Elementary Kinetic Theory of the Viscosity of a Gas. One of the most interesting applications of mechanics to kinetic theory is provided by the theory of gas viscosity. From elementary physics the reader will recall that in all real fluids in motion

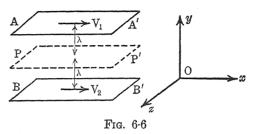
¹ For further discussion, cf. Lindsay, op. cit., pp. 76 ff, 218, 234, 236.

whether liquid or gaseous, there is always present a tendency for the fluid moving in one layer to retard the motion of the fluid in the immediately adjacent layers. This effect is attributed to a force of viscosity or viscous drag, which Newton assumed is directly proportional to the flow velocity gradient or rate of change of velocity with distance perpendicular to the direction of flow. It is moreover assumed that the viscous force is also proportional to the area of contiguous layers. If we denote the area in question by A, the flow velocity gradient by dV/ds, the viscous drag F_v may then be written

$$F_v = \eta A \frac{dV}{ds}, \qquad (6.9-1)$$

where η is the coefficient of viscosity or, more briefly, the *viscosity* of the fluid. It is the tangential viscous drag per unit area per unit flow velocity gradient. In absolute units its dimensions are dyne second per cm². For example, the value for water in these units at 20° C is 0.01, while for hydrogen at 0° C η is 8.4 × 10⁻⁵. It is of interest to note that the viscosity of liquids decreases as the temperature rises, whereas that of gases increases with the temperature.

The viscosity of liquids finds a plausible explanation in the cohesive forces between the constituent parts, but this is not available for an ideal gas in which such forces are assumed not to exist or are at any rate ignored. Maxwell was the first to give a kinetic theory explanation of gas viscosity in terms of the motion of the molecules and in particular by the transfer of momentum from one moving layer of gas to another by the random molecular motion. Let us look into this matter in an elementary fashion. Consider a gas which has a flow movement from left to right, let us say in the x direction. Let the velocity in the plane AA' (Fig. 6.6) be denoted by V_1 , while that in the lower plane BB' at a distance 2λ from AA' is $V_2 < V_1$. Draw the imaginary plane PP'half way between AA' and BB'. Now we may consider the molecules moving in the plane to be divided into three groups, namely, those traveling in the direction of the x, y and z axes respectively. We may suppose that if there are n molecules per unit volume, there will be on the average n/3 proceeding in each of the three directions. Hence on the average the number of molecules that travel across unit area of the plane PP' in the positive y direction per second is $n/6 \cdot v_m$, where v_m is the root mean square molecular velocity, and we are making the simplifying assumption that all the molecules may be considered as moving with this velocity. The same number on the average cross unit area of the plane in the negative y direction per second. By virtue of the way



in which the planes AA' and BB' are drawn, the above-mentioned molecules have each suffered their last collision (before striking PP') in either the plane AA' or BB', since λ is by definition the mean free path. Now the velocity gradient of the gas in the neighborhood of PP' is approximately

$$\frac{dV}{ds} = \frac{V_1 - V_2}{2\lambda} \tag{6.9-2}$$

It is, of course, assumed that the velocities V_1 and V_2 are very small compared with the molecular velocity v_m . We shall next suppose that when a molecule passes through any plane it instantaneously acquires the velocity with which the gas as a whole is moving in that plane, the latter then being compounded with the molecular velocity. Hence we can say that the $nv_m/6$ molecules which travel across unit area of the plane PP' per second coming from plane AA' carry with them or transfer a certain amount of momentum. Each such molecule has acquired what may be called a "flow" momentum (to distinguish it from the momentum due to its velocity v_m) equal to mV_1 . Hence in the passage of the molecules from AA' through PP' there is a transfer of momentum per second per unit area from the gas above PP' to the gas below PP' amounting to

$$\frac{nv_m}{6} \cdot mV_1. \tag{6.9-3}$$

But we know that a change in momentum implies a force which

indeed is equal to the time rate of change of momentum. Hence we may look upon the expression (6.9-3) as representing a tangential force exerted by the gas above a unit area of the plane PP' on the gas below (tangential because the change in momentum is in the direction in which the gas flow is taking place). At the same time the passage of $nv_m/6$ molecules per second from the layer beneath PP' transfers in the upward direction tangential momentum per second per unit area to the amount of

$$\frac{nv_m}{6} \cdot mV_2. \tag{6.9-4}$$

This represents the tangential stress exerted by the gas below the plane PP' on the gas above. Now by Newton's third law (Sec. 1·7) the equal and opposite reaction on the gas below is $-nv_m/6 \cdot mV_2$. Hence the resultant tangential drag on unit area of the gas immediately below PP' is given by

$$\frac{nv_m}{6} \cdot m(V_1 - V_2). \tag{6.9-5}$$

By definition this is F_v/A in (6.9-1). From this and (6.9-2) the viscosity of the gas comes out to be

$$\eta = \frac{\rho v_m \lambda}{3} \cdot \tag{6.9-6}$$

The deduction of this fundamental relationship given above is somewhat crude and the standard kinetic theory texts¹ should be consulted for a more elaborate treatment. Our purpose has been to bring out the fundamental rôle played by the transfer of momentum. If we employ (6.7-9) and (6.8-11) the relation (6.9-6) takes the form

$$\eta = \frac{\sqrt{3kmT}}{4\pi D^2}. (6.9-7)$$

This indicates that the viscosity of an ideal gas is independent of the pressure and hence also of the density, a prediction first made by Maxwell and experimentally verified by him over a wide range

¹ Cf. Loeb, op. cit., p. 180 ff. Also E. H. Kennard, "Kinetic Theory of Gases" (McGraw-Hill, New York, 1938) p. 138 ff.

of values.¹ The temperature dependence has also received ample verification. It is of value to emphasize that the above result in no way depends on any assumed forces of interaction between molecules.

A number of striking numerical results can be readily obtained from the equations of this section and the preceding one. Thus using the experimental value of η for hydrogen at 0° C. as 8.41 \times 10⁻⁵ gm/cm sec we can calculate the mean free path of the hydrogen molecule at 0° C and atmospheric pressure. We have

$$\lambda = \frac{3\eta}{\rho v_m} = \frac{3 \times 8.41 \times 10^{-5}}{8.99 \times 10^{-5} \times 1.83 \times 10^{5}} \text{ cm},$$
$$= 1.53 \times 10^{-5} \text{ cm}.$$

This gives one a good idea of the order of magnitude of λ . From this the average number of collisions experienced by a hydrogen molecule per second can at once be obtained. Thus

$$N_c = \frac{1.83 \times 10^5}{1.53 \times 10^{-5}} \,\text{sec}^{-1} = 1.20 \times 10^{10} \,\text{sec}^{-1}.$$

Lastly we may obtain some conception of the size of a molecule from eq. (6.8-11). For

$$D = \left(\frac{0.75}{n\pi\lambda}\right)^{\frac{1}{2}},\tag{6.9-8}$$

and taking $n=2.705\times 10^{19}$ as the number of molecules per cm³ (by Avogadro's hypothesis this is the same for all gases at the same temperature and pressure),² and using the given value above of λ we have approximately

$$D = 2.4 \times 10^{-8} \text{ cm}.$$

There have been many estimates of the size of molecules on the kinetic theory, as the reader will find by an examination of texts on this subject. The interesting fact is that they all agree, so far as order of magnitude is concerned, with the foregoing result, which thus seems to have fundamental significance. Of course,

¹ At very high densities (e.g. 100 atmospheres) this result no longer follows and the viscosity increases more or less in proportion to the density. A similar result obtains at very low densities. (Cf. Kennard, op. cit., p. 150).

² R. T. Birge, "Probable Values of the General Physics Constants."

we must be careful not to think of this necessarily as the diameter of a solid sphere. Modern physics has a great deal to say about the structure of molecules and the atoms which are their constituents. We have already had occasion to note the Bohr-Rutherford theory of atomic structure. On this point of view the atom and molecule have very *open* structures, so that the quantity represented by D will here presumably have some connection with the spatial separation of the component parts.

6·10. Virial for Interaction Forces. Kinetic Theory of a Real Gas. In our discussion of the kinetic theory of Secs. 6·7 and 6·8 we have confined our attention entirely to an ideal gas of freely moving particles with no interaction forces between them. The discrepancy between the results obtained (e.g. the equation of state) and experimental formulas has been largely attributed to the existence of just such forces. It may be worth while to note their contribution to the virial of a collection of particles. Reverting to the general definition of the virial in (6·6-7), if \mathbf{F}_i has the form (6·5-4), we get for N particles

$$\Omega = -\frac{1}{2} \sum_{j,k=1}^{N} f(r_{jk}) \mathbf{r}_{jk} \cdot \mathbf{r}_{j}, \qquad (6.10-1)$$

where both j and k are summed from 1 to N but the terms corresponding to j=k are, of course, omitted. The sum consists of pairs of terms of the form

$$f(r_{jk})(\mathbf{r}_j - \mathbf{r}_k) \cdot \mathbf{r}_j + f(r_{kj})(\mathbf{r}_k - \mathbf{r}_j) \cdot \mathbf{r}_k$$

= $f(r_{jk})r_{jk}^2$. (6·10-2)

Hence the contribution of these forces to the virial becomes

$$\Omega = -\frac{1}{2} \sum_{j < k=1}^{N} f(r_{jk}) r_{jk}^{2}.$$
 (6·10-3)

Consequently the complete equation of state will not be (6.7-11) but rather

$$pV_0 = NkT + \frac{1}{3} \sum_{j,k=1}^{N} f(r_{jk}) r_{jk}^2.$$
 (6·10-4)

If we knew the precise character of the interaction force, i.e., the function $f(r_{jk})$, and could evaluate the sum in the second term on

the right in $(6\cdot10-4)$, we should possess a theoretical deduction of the equation of state of a real gas. However, we do not know the $f(r_{jk})$ precisely and even if we make various plausible assumptions the mathematical evaluation of the sum is practically impossible in anything like simple analytical terms. Certain approximate methods lead from $(6\cdot10-4)$ to the famous Van der Waals' equation of state

$$\left(p + \frac{a}{V_0^2}\right) (V_0 - b) = NkT,$$
 (6·10-5)

but for the discussion of this deduction the reader is referred to the standard treatises.¹

PROBLEMS

- 1. An automobile of mass 2 tons traveling due south at 50 mi/hr collides at an intersection with a 5-ton truck traveling due east at 30 mi/hr. Assuming that the two cars lock together after impact, what is the velocity in magnitude and direction after collision? What is the velocity (magnitude and direction) of the center of mass of the two cars before and after collision? Calculate the total kinetic energy of the cars before and after collision and comment on the physical meaning of the result.
- 2. An alpha particle is projected from position x_1 in the positive x direction with speed v_1 . At the same instant another alpha particle is projected from position x_2 ($x_2 > x_1$) in the negative x direction with speed v_2 . Describe the subsequent motion of the two particles.
- 3. Two particles of masses m_1 and m_2 respectively move in arbitrary fashion. Prove that the total moment of momentum of the system about the center of mass is a vector at right angles to the plane containing the line joining the instantaneous positions of the particles and the instantaneous relative velocity \mathbf{v}_r of the two particles and has the magnitude

$$\left(\frac{m_1m_2}{m_1+m_2}\right)\cdot pv_r,$$

where p is the perpendicular distance from one of the particles to the line through the other particle having the direction of \mathbf{v}_r .

- 4. The two particles in an Atwood's machine have masses m_1 and m_2 grams respectively and the two strings are separated by a distance r. If the system is allowed to move from rest when the two masses are at the same height h from the floor, find the position of the center of mass after t seconds. Also find the kinetic energy and potential energy (with respect to the floor) of each particle at the end of the t seconds.
 - ¹ A simplified version is found in Lindsay, "Physical Statistics," p. 96 ff.

- 5. A shell is observed to explode when at the highest point of its path. It splits into two parts of equal mass of which one is observed to fall vertically. Find the path followed by the other with respect to the ground.
- 6. Two particles of masses m_1 and m_2 respectively move in arbitrary fashion. If \mathbf{v}_c is the velocity of the center of mass and \mathbf{v} the relative velocity of the two particles with respect to each other show that the total kinetic energy of the system is

$$\frac{1}{2} (m_1 + m_2) v_c^2 + \frac{1}{2} \left(\frac{m_1 m_2}{m_1 + m_2} \right) v^2.$$

- 7. Calculate the position of the center of mass of the system composed of the sun and the planet Jupiter. Compare the kinetic energy of the system at perihelion with that at aphelion.
- 8. Two particles of masses m_1 and m_2 respectively move about each other. Calculate the exact expression for the periodic time, i.e., the time for one complete period of the motion.
- 9. Two particles which move subject only to their mutual attraction are projected in any direction with any velocities. Show that the line joining the particles always remains parallel to a fixed plane.
- 10. A uniform chain 2 meters long hangs over a perfectly smooth peg with initially 110 cm on one side and 90 cm on the other. How long will it take the chain to slide off the peg and what velocity will it have at that instant?
- 11. A sphere of radius R contains an ideal gas. Derive the expression for the virial due to the impact of the molecules on the surface of the sphere.
- 12. In the Bohr model of a single electron atom (see 3.9) the electron is assumed to possess a quantized moment of momentum (also called angular momentum) given by

$$p_{\theta} = n_{\theta} h / 2\pi,$$

where n_{θ} is an integer. Suppose that the electron also possesses a spin about an axis through itself. Let this spin correspond to another quantized moment of momentum

$$p_s = n_s h/2\pi,$$

where n_s is integral. Recalling that the moment of momentum is a vector quantity, find the magnitude of the resultant moment of momentum of a spinning electron revolving in a quantized orbit if the two moments make angle θ with each other. If the resultant moment of momentum is restricted in magnitude to $nh/2\pi$ where n is half-integral, show that θ may not take on all values. Find the allowed quantized values of θ and specialize to the case where $n_s = \pm \frac{1}{2}$, $n_\theta = 1$ and $n = \frac{3}{2}$.

13. A stationary nucleus of mass m is struck head on by a neutron of mass m_n moving with velocity v_0 . If the collision is perfectly elastic, find the expression for the recoil velocity of the nucleus. How could this equation be used to determine the mass of the neutron, even if the velocity v_0 is not known?

174 MOTION OF A SYSTEM OF PARTICLES

- 14. Compare the root mean square velocity of hydrogen molecules at temperature 300° K with that at liquid air temperature and at the temperature of the sun. Carry out similar calculations for a number of other gases. Indicate the reason why there is little or no hydrogen in the earth's atmosphere.
- 15. Determine the dependence of the mean free path of an ideal gas on the pressure and temperature. Calculate the mean free path in hydrogen at 100° C and a pressure of 10^{-6} atmosphere.
- 16. The mass of an electron is 9×10^{-28} gram. Assuming a gas composed entirely of electrons, calculate their root mean square velocity at $T=300^{\circ}$ K. Compare the pressure they would exert on the walls of a container with the pressure that hydrogen molecules at this temperature would exert on the walls of an identical container, if the concentration of particles is the same in both cases.
- 17. Calculate the value of R, the gas constant per gram molecule. Use this to compute the specific heat at constant volume for a number of monatomic gases (e.g., rare gases).
- 18. Derive the expression for the quantized energy of the two-particle system composed of an electron and a nucleus revolving about their common center of mass (cf. Sec. 4.4 for the equivalent one-particle problem).





CHAPTER VII

MECHANICS OF A RIGID BODY

7.1. Definition of a Rigid Body. Types of Motion. Chapter VI was concerned with the motion of an aggregate of material particles. An important special case is that in which the forces of interaction of the particles are such that the distance between any two particles remains constant in time. The aggregate is then called a *rigid body*. Since the interaction forces in actual solids, though strong, are finite and all such solids can be deformed by sufficiently strong external forces, it is clear that there exist no really rigid bodies. Nevertheless, many bodies act sufficiently like the ideal rigid body just defined to make the study of their motions and general behavior a significant branch of mechanics.

The general theorems established in the preceding chapter may be applied at once to rigid bodies. But certain simplifications are apparent since the possible types of motion are restricted. If one point of the body is fixed with respect to the primary inertial system, the only possible motion is that in which every other point moves on the surface of a sphere whose radius is the invariable distance from the moving point to the fixed point. If two points of the body are fixed, the only possible motion is that in which all points save those on the line joining the two fixed points (extended, of course) move in circles about centers located on the line. Finally if three points of the body not in the same straight line are fixed, it is unable to move at all, and its position is completely determined.

Motion of a rigid body in which one point is kept fixed or two points are kept fixed is known as motion of *rotation*. This is simpler than the general motion of the body when no point is kept fixed. Nevertheless, there is a special case of the latter which is elementary, namely, that in which all points of the body move in the same direction at any instant with the same velocity and acceleration. This is called motion of *translation*. The reader is advised to take a meter stick or similar object and exemplify to himself these two types of motion. They are the most important

types of motion of a rigid body, since it can be shown that every displacement of a rigid body can be considered as a combination of translations and rotations. Consider Fig. 7-1 and suppose the problem is to get the rigid body shown schematically there from position I to position III in which three points of the body which were originally at positions A, B, C now appear at A', B', C'. First translate the body so that A moves to A', its new desired position.

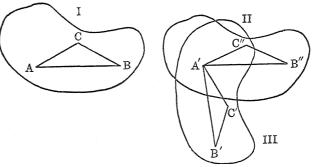


Fig. 7.1

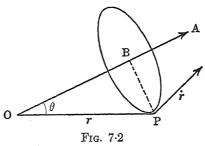
Then from the definition of translation, AB will become A'B'', where A'B'' is parallel to AB, and likewise A'C'' is parallel to AC, and B''C'' is parallel to BC. Now holding A' fixed, rotate the body so that B'' coincides with B', the desired final position of B. Without moving A' or B' it will now be possible to make C'' coincide with C', the desired final position of C, since the triangle A'B'C' is congruent to the triangle A'B''C''. Hence it has been possible to go from the original to the final position by a translation followed by a motion in which one point remains fixed, i.e. a rotation.

The translation of a rigid body will be given by the translation of any point in it, e.g., the center of mass (Sec. 6·1). Hence the mechanics of particle motion will take care of this, and we need not consider it further. In what follows we shall concentrate on rotation.

7.2. Rotation and Angular Velocity. Let us consider an axis OA fixed in space (cf. Fig. 7.2) either in the rigid body or outside it. From the definition given in the preceding section, if the body rotates about OA, it means that any point of the body such as P

moves in a circle with center B on the line OA (extended if necessary), where B is the normal projection of P on OA. Let r be the

position vector of P relative to O as origin. Then the circle of motion of P lies in a plane perpendicular to the plane of r and the axis OA. The velocity of P, namely, \dot{r} , always lies in the plane of the circle and hence is perpendicular to the plane AOP. In time dt, P moves through the arc



 $|\dot{\mathbf{r}}| dt$. But this equals the radius $BP(=r\sin\theta)$ multiplied by the angle through which BP moves. Call the latter $d\phi$. We then have

$$|\dot{\mathbf{r}}| dt = r \sin \theta \cdot d\phi, \tag{7.2-1}$$

or

$$|\dot{\mathbf{r}}| = r \sin \theta \cdot \frac{d\phi}{dt} = r\omega \sin \theta,$$
 (7.2-2)

where we have written ω for $d\phi/dt$. If we revert to Sec. 6.4 (eq. 6.4–1) we see that it is reasonable to interpret $r\omega \sin \theta$ in eq. (7.2–2) as the magnitude of the cross product of the position vector \mathbf{r} with a new vector of magnitude ω whose direction makes the angle θ with \mathbf{r} , in such a way that $\dot{\mathbf{r}}$ is perpendicular to the plane of the new vector and \mathbf{r} . If the new vector lies along OA and is directed from O to A, it will fulfill these requirements precisely. We shall call it the instantaneous angular velocity of the rigid body (it is the same for every particle by virtue of the definition of rigid body) and represent it by ω . The vector formulation corresponding to (7.2–2) then becomes

$$\dot{\mathbf{r}} = \mathbf{\omega} \times \mathbf{r}. \tag{7.2-3}$$

The connection between the definition of angular velocity ω for a rigid body as given here and the definition of angular speed of a particle moving in a circle as given in Sec. 1.4 should be carefully noted.

It can be shown that ω as defined above possesses the usual properties of vectors (cf. Sec. 1.3). In particular the resultant of two angular velocities passing through the same point is found by

using the customary parallelogram rule. Hence we can always resolve any angular velocity ω into rectangular components ω_z , ω_y , ω_z such that

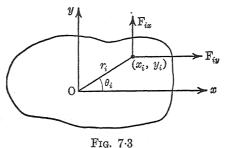
$$\omega = i\omega_x + j\omega_y + k\omega_z. \tag{7.2-4}$$

It is well to emphasize that the vector property of angular velocity is not shared by a finite rotation, even though the latter can be given a direction and a magnitude. It is not therefore customary to consider a finite rotation as representable by a vector.

The time rate of change of angular velocity is termed the angular acceleration. Thus

$$\mathbf{a} = \frac{d\mathbf{\omega}}{dt} \cdot \tag{7.2-5}$$

7.3. Rotation about a Fixed Axis. The simplest case of rotation of a rigid body is that of rotation about an axis fixed in space. We suppose the axis chosen as the z axis of an appropriate system of



coördinates and indicate the trace of the rigid body in the xy plane in Fig. 7.3. Consider the mass particle m_i in this plane with coördinates x_i , y_i (such that $r_i^2 = x_i^2 + y_i^2$). Suppose the force F_i acts on the particle in the xy plane and let its x and y components respectively be F_{ix} and F_{iy} . These produce a force moment or torque (cf. Sec. 6.4) about the z axis with magnitude equal to

$$L_{i} = x_{i}F_{iy} - y_{i}F_{ix} = m_{i}x_{i}\ddot{y}_{i} - m_{i}y_{i}\ddot{x}_{i} = \frac{d}{dt}[m_{i}(x_{i}\dot{y}_{i} - y_{i}\dot{x}_{i})].$$
(7.3-1)

Now from the figure we have

$$x_i = r_i \cos \theta_i, \quad y_i = r_i \sin \theta_i,$$
 (7.3-2)

whence

$$\dot{x}_i = -r_i \omega \sin \theta_i = -\omega y_i, \quad \dot{y}_i = r_i \omega \cos \theta_i = \omega x_i. \quad (7.3-3)$$

Note that $\omega = \dot{\theta}_i$ is independent of *i* because of the rigidity. If we substitute from (7·3-3) into (7·3-1) we get

$$L_{i} = \frac{d}{dt} \left[\omega m_{i} (x_{i}^{2} + y_{i}^{2}) \right] = \frac{d}{dt} \left[\omega m_{i} r_{i}^{2} \right]. \tag{7.3-4}$$

Let us now sum over all the particles of the body and arrive at the magnitude of total torque about the z axis. We have a right to sum because all the individual torques are vectors directed along the z axis. Thus

$$L = \sum L_i = \frac{d}{dt} \left[\omega \sum_i m_i r_i^2 \right]. \tag{7.3-5}$$

Note the quantity $\sum_{i} m_{i} r_{i}^{2}$ which is formed by multiplying the mass of each particle by the square of its distance to the z axis or axis of rotation. We call this the *moment* of *inertia* of the rigid body about the z axis and denote it here simply by I. Eq. (7.3-5) then becomes

$$L = \frac{d}{dt} (I\omega). \tag{7.3-6}$$

The formal analogy between this and Newton's second law (1.7-5) is very striking, and it is not surprising that $I\omega$ is known as the angular momentum about the fixed axis. As presented here for the special case of rotation about a fixed axis it is a scalar quantity. We shall generalize the definition somewhat later.

If the body is really rigid, I is independent of time and (7.3–6) becomes

$$L = I\alpha, \tag{7.3-7}$$

whose analogy with F = ma is worthy of notice.

Still restricting ourselves to the fixed axis, let us calculate the work done by the forces \mathbf{F}_i on the rigid body. If the particle at (x_i, y_i) makes a displacement with components dx_i and dy_i , respectively, the work done is clearly

$$dW_{i} = F_{ix} dx_{i} + F_{iy} dy_{i}. (7.3-8)$$

But we have from (7.3-2)

$$dx_{i} = -r_{i} \sin \theta_{i} d\theta = -y_{i} d\theta,$$

$$dy_{i} = r_{i} \cos \theta_{i} d\theta = x_{i} d\theta.$$

$$(7.3-9)$$

Here we replace $d\theta_i$ by $d\theta$ since, when the rigid body rotates about the fixed axis, all particles move through the same angle for a given displacement. Substituting into (7.3-8) yields

$$dW_i = (x_i F_{iy} - y_i F_{ix}) d\theta = L_i d\theta. \tag{7.3-10}$$

Hence the work done on the whole body during a rotation from θ_0 to θ_1 becomes

$$W = \int_{\theta_0}^{\theta_1} \sum L_i d\theta = \int_{\theta_0}^{\theta_1} \frac{d}{dt} (I\omega) d\theta.$$
 (7.3–11)

If I is constant, this results in

$$W = \int_{\omega_0}^{\omega_1} I\omega \, d\omega = \frac{1}{2} I\omega_1^2 - \frac{1}{2} I\omega_0^2, \qquad (7.3-12)$$

where ω_0 is the initial angular velocity magnitude, i.e., that corresponding to the initial angle θ_0 and ω_1 is the final angular velocity component. The analogy between $(7\cdot3-12)$ and the work-kinetic energy theorem for the motion of a particle (eq. 1·10-20) is clear. We naturally refer to $I\omega^2/2$ as the kinetic energy of rotation, and eq. $(7\cdot3-12)$ is called the work-kinetic energy theorem for rotation. The units of $I\omega^2/2$ are the same as those of $mv^2/2$.

7.4. Moment of Inertia Calculations. Since the moment of inertia is such an important quantity in the motion of rigid bodies we ought to devote some attention to its calculation in special cases. As we have already seen, if we consider the body in question to be a collection of discrete particles of masses m_1, m_2, \ldots, m_n , rigidly connected, the moment of inertia about the given axis of rotation is

$$I = \sum_{i=1}^{n} m_i r_i^2, (7.4-1)$$

where r_i is the perpendicular distance from the mass m_i to the axis. In most cases arising in practice, however, the body is a continuous distribution of mass and for the actual calculation of I the summa-

tion in (7·4-1) must be replaced by the corresponding integral extended over the whole body, viz.:

$$I = \int r^2 \, dm = \int \rho r^2 \, dV, \qquad (7.4-2)$$

where dV is an appropriately chosen volume element, ρ is the value of the density of the substance composing the body at the element, and r is the perpendicular distance from the element to the axis. The determination of I then reduces to the mathematical problem of evaluating a definite volume integral. The ease with which this is accomplished in a given case depends largely on the appropriate choice of dm or dV.

Let us take first a very simple illustration. Find the moment of inertia of a homogeneous solid rod of constant linear density ρ_l (grams/cm, say) and of length l about an axis perpendicular to the

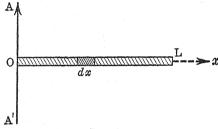


Fig. 7.4

rod through one end (Fig. 7-4). In the figure, OL represents the rod and AOA' the axis. We choose the x axis of a rectangular system along OL extended with the origin at O. The mass element here is clearly

$$dm = \rho_l dx,$$

and hence

$$I = \int_0^l \rho_l x^2 \, dx = m_l \, \frac{l^2}{3}, \qquad (7.4-3)$$

where ρ_l l has been replaced by m_l , the mass of the rod.

There is a simple but important interpretation of the result embodied in (7.4-3). We note at once that the behavior of the rod, in so far as its rotation about the axis AOA' is concerned, may be studied by replacing the rod by a particle of mass $m_l/3$ whose distance from the axis is l, or by a particle of mass m_l whose distance

from the axis is $l/\sqrt{3}$. Ordinarily greater significance is attributed to the latter type of replacement. The distance $l/\sqrt{3}$ is called the radius of gyration of the rod with respect to the axis of rotation. We may immediately generalize this definition by setting in all cases

 $k = \sqrt{\frac{I}{m}}, \qquad (7.4-4)$

where k is the radius of gyration of the rigid body of mass m, whose moment of inertia about the axis in question is I.

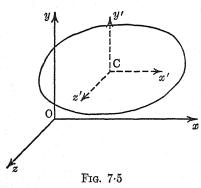
Coming back to the solid rod, if we were to calculate the value of I about a perpendicular axis through the center of the rod (i.e., in this case the center of mass), we should have

$$I_C = \int_{-l/2}^{+l/2} \rho_l \, x^2 \, dx = m_l \, \frac{l^2}{12} \cdot \tag{7.4-5}$$

It is seen that we have the relation

$$I = I_C + m_l \left(\frac{l}{2}\right)^2 \cdot \tag{7.4-6}$$

This is a special case of a general law called the theorem of parallel axes. We may state it as follows: The moment of inertia of a rigid body about any axis is equal to the moment of inertia about a



parallel axis through the center of mass plus the product of the mass of the body and the square of the perpendicular distance between the two axes. Let us prove the theorem for the general case. Consulting Fig. 7.5, we shall take the z axis as the axis of rotation and let C with coördinates \overline{x} , \overline{y} , \overline{z} be the center of mass. Treating the body as a set of mass particles, let us suppose that a

particle m_i with coördinates x_i , y_i , z_i in the original system has the coördinates x_i' , y_i' , z_i' when referred to a set of parallel axes through C. Now since $\bar{x}' = \bar{y}' = \bar{z}' = 0$, we must have

$$\sum m_i x_i' = \sum m_i y_i' = \sum m_i z_i' = 0, (7.4-7)$$

from the definition of center of mass (Sec. 6-1). By definition the moment of inertia about the z axis is

$$I = \sum m_i (x_i^2 + y_i^2). \tag{7.4-8}$$

But

$$x_i = x_i' + \overline{x}, \quad y_i = y_i' + \overline{y},$$

and hence

$$I = \sum m_i (x_i'^2 + y_i'^2) + (\bar{x}^2 + \bar{y}^2) \sum m_i + 2\bar{x} \sum m_i x_i' + 2\bar{y} \sum m_i y_i'.$$
(7.4-9)

Now the last two terms in $(7\cdot4-9)$ vanish by virtue of $(7\cdot4-7)$. Then $\sum m_i$ is the total mass of the body, while $\bar{x}^2 + \bar{y}^2 = l^2$, where l is the perpendicular distance between the z and z' axés. $\sum m_i(x_i'^2 + y_i'^2)$ is the moment of inertia with respect to the z' axis (i.e., axis through center of mass parallel to the z axis). Hence in general

$$I = I_c + ml^2. (7.4-10)$$

To continue our brief discussion of the calculation of I in special cases, let us compute the moment of inertia of a homogeneous

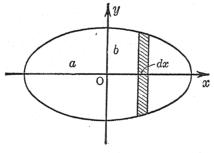


Fig. 7.6

elliptical disk of mass m, superficial density σ , and semi-axes a and b, about the two axes respectively (Fig. 7.6).

If the x and y axes are chosen along the major and minor axes respectively the equation of the ellipse is

$$\frac{x_{\star}^2}{a^2} + \frac{y^2}{b^2} = 1. (7.4-11)$$

First let us calculate I about the minor axis. Taking a strip of

width dx at distance x from the origin, we find its mass is $2\sigma y dx$ and its moment of inertia about the y axis

$$2\sigma yx^2 dx$$
.

Hence the total moment for the whole disk is because of symmetry

$$I_y = 4\sigma \int_0^a x^2 y \, dx = \frac{4\sigma b}{a} \int_0^a x^2 \sqrt{a^2 - x^2} \, dx, \quad (7.4-12)$$

utilizing eq. (7·4-11). From Peirce's tables (No. 145) the integration is readily carried out, yielding

$$I_y = \pi \sigma b \frac{a^2}{4} = m \frac{a^2}{4},$$
 (7.4-13)

where we have placed $m = \sigma \pi ab$, the area of the ellipse being πab . Similarly the moment of inertia about the major axis is

$$I_x = m \frac{b^2}{4} \cdot \tag{7.4-14}$$

As a final illustration let us take a solid body, say a sphere of radius a and calculate the moment of inertia about a diameter.

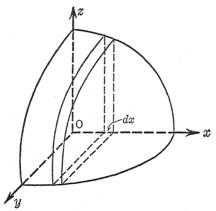


Fig. 7.7

The equation of the sphere with center at the origin is $x^2 + y^2 + z^2 = a^2$. Consider the octant shown in Fig. 7.7 and cut out the slice of thickness dx at distance x from the yz plane. If now we can find the moment of inertia of this slice or rather the whole

circular disk of which this is one quadrant, about the x axis (i.e., perpendicular axis through the center), we can calculate the total moment by integration with respect to x with limits -a and +a.

Let us consider then the case of the circular disk of radius R. As mass element (Fig. 7.8) take the ring of radius r and thickness dr. The superficial density being σ , we have

$$I = 2\pi\sigma \int_0^R r^3 dr = 2\pi\sigma \frac{R^4}{4}$$
$$= \frac{1}{2}mR^2, \qquad (7.4-15)$$

if m is here the mass of the disk.

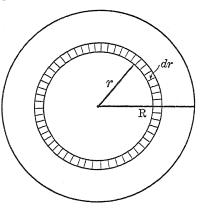


Fig. 7.8

We now utilize the foregoing

result in the problem of the sphere. The radius of the disk in question is $\sqrt{a^2-x^2}$ and its mass is $\rho dx \cdot \pi (a^2-x^2)$, so that the moment about the x axis is from $(7\cdot 4-15)$

$$\frac{1}{2} \cdot \pi \rho (a^2 - x^2)^2 dx.$$

Hence the total moment is

$$I = \pi \rho \int_0^a (a^2 - x^2)^2 dx$$
$$= \pi \rho \frac{8a^5}{15}.$$
 (7.4–16)

Now the mass of the sphere is

$$m=\tfrac{4}{3}\pi a^3\rho,$$

and hence we may write

$$I = \frac{2}{5}ma^2. (7.4-17)$$

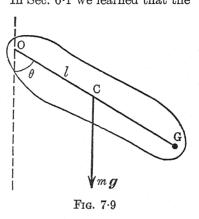
The following table includes the values of the moment of inertia for a number of special cases which are likely to prove useful to the student.

Moment of Inertia of Various Bodies $m = {\rm mass~of~the~body}$

Body	Axis	Moment of Inertia
Thin rectangular sheet of sides a and b	Through the center paral- lel to b	$m\frac{a^2}{12}$
Thin rectangular sheet of sides a and b	Through the center perpendicular to the sheet	$m\frac{a^2+b^2}{12}$
Thin circular sheet of radius r	Normal to the plate through the center	$m\frac{r^2}{2}$
Thin circular sheet of radius r	Along any diameter	$m\frac{r^2}{4}$
Thin circular ring. Radii r_1 and r_2	Through center normal to plane of ring	$m \frac{r_1^2 + r_2^2}{2}$
Thin circular ring. Radii r_1 and r_2	A diameter	$m \frac{r_1^2 + r_2^2}{4}$
Spherical shell, very thin, mean radius, r	A diameter	$m\frac{2r^2}{3}$
Thin cylindrical shell, radius r , length l	Longitudinal axis	mr^2
Right circular cylinder of radius r , length l	Longitudinal axis	$m\frac{r^2}{2}$
Right cone, altitude h , radius of base r	Axis of the figure	$m \frac{3}{10} r^2$
Spheroid of revolution, equatorial radius r	Polar axis	$m\frac{2r^2}{5}$
Ellipsoid, axes $2a$, $2b$, $2c$	Axis 2a	$m\frac{(b^2+c^2)}{5}$
Uniform thin rod	Normal to the length, at one end	$m \frac{l^2}{3}$
Uniform thin rod	Normal to the length, at the center	$m \frac{l^2}{12}$
Rectangular prism, dimensions 2a, 2b, 2c	Axis $2a$	$m\frac{(b^2+c^2)}{3}$
Sphere, radius r	A diameter	$mrac{2}{5}r^2$
Rectangular parallelepiped, edges a , b , and c	Through center perpendicular to face ab (parallel to edge c)	$m\frac{a^2+b^2}{12}$
Right circular cylinder of radius r , length l	Through center perpen- dicular to the axis of the figure	$m\left(\frac{r^2}{4} + \frac{l^2}{12}\right)$
Spherical shell, external radius r_1 , internal radius r_2	A diameter	$m \frac{2}{5} \frac{(r_1^5 - r_2^5)}{(r_1^3 - r_2^3)}$
Hollow circular cylinder, length l , external radius r_1 , internal radius r_2	Longitudinal axis	$m \frac{(r_1^2 + r_2^2)}{2}$
Hollow circular cylinder, length l , radii r_1 and r_2	Transverse diameter	$m\left[\frac{r_1^2 + r_2^2}{4} + \frac{l^2}{12}\right]$
Hollow circular cylinder, length l , very thin, mean radius r	Transverse diameter	$m\left(\frac{r^2}{2} + \frac{l^2}{12}\right)$

7.5. The Physical Pendulum. An interesting practical illustration of the motion of a rigid body about a fixed axis is provided by the oscillations due to gravity of a body suspended from a fixed horizontal axis. This is the so-called physical or compound pendulum (Fig. 7.9). In order to apply the analysis of Sec. 7.3 to this problem, it is necessary to calculate the total torque about the axis of suspension through O. In Sec. 6.1 we learned that the

center of mass of a collection of particles is the point which moves as if all the mass were concentrated there and as if all the external forces acted there. In the present problem the external force acting is the force of gravity on every particle of the body. To think of all these forces acting at the center of mass is to suppose that the whole weight of the body mg acts at the center of mass. This enables us at once to calculate the resultant torque about O.



Assume that the center of mass (cf. Sec. 7.8 for calculations of the position of the center of mass of continuous aggregates) is at C, distant l from the axis of suspension.

The resultant torque about the axis for any displacement θ has therefore the magnitude

$$L = mgl \sin \theta. \tag{7.5-1}$$

If we denote the radius of gyration about the center of mass by k, the moment of inertia about the axis of suspension by the parallel axis theorem is $I = m(k^2 + l^2)$. Hence the fundamental eq. (7.3-6) takes the form

$$-mgl\sin\theta = m(k^2 + l^2)\ddot{\theta},$$

or

$$\ddot{\theta} = -\frac{gl\sin\theta}{k^2 + l^2}. (7.5-2)$$

Now for small displacements $\sin \theta \doteq \theta$, and hence if we confine our-

selves to such, we have

$$\ddot{\theta} = -\frac{gl\theta}{k^2 + l^2} \tag{7.5-3}$$

We have met this equation before (cf. Sec. 2.2). It is indeed the equation of simple harmonic angular motion and the solution is

$$\theta = A \sin \left(\sqrt{\frac{gl}{k^2 + l^2}} t + B \right), \tag{7.5-4}$$

so that the frequency of the motion is

$$\nu = \frac{1}{2\pi} \sqrt{\frac{gl}{k^2 + l^2}},\tag{7.5-5}$$

or expressed directly in terms of the moment of inertia I

$$\nu = \frac{1}{2\pi} \sqrt{\frac{mgl}{I}}$$
 (7.5–6)

The amplitude A is arbitrary within the limits of the assumption of small displacements. The arbitrary constant B is the initial phase. Both A and B must be determined by the initial conditions of the motion.

A special case of significance is that in which the whole mass of the body is concentrated at the point C, and the rest of the body is replaced by a theoretically massless cord of length l_1 connecting C to the point of suspension. Then $I = ml_1^2$ at once, and we have

$$\ddot{\theta} = -\frac{g\sin\theta}{l_1},\tag{7.5-7}$$

which under the same approximation as before yields

$$\theta = A \frac{\sin}{\cos} \left(\sqrt{\frac{g}{l_1}} t + B \right), \tag{7.5-8}$$

where the frequency and period of the motion are respectively

$$\nu = \frac{1}{2\pi} \sqrt{\frac{g}{l_1}},$$
and
$$P = 2\pi \sqrt{\frac{l_1}{g}}.$$
(7.5–9)

and

Such a pendulum is called a *simple pendulum*. From (7.5-5) and (7.5-9) we see that the period of the *physical* pendulum with moment of inertia I and distance from the center of mass to the point of suspension l, is equal to the period of a *simple* pendulum of length l_1 , where

$$l_1 = \frac{I}{ml} = \frac{k^2 + l^2}{l} {.} (7.5-10)$$

The length l_1 satisfying this condition is called the length of the equivalent simple pendulum. This means that if we extend the line OC in Fig. 7.9 to a point G on the other side of C such that $OG = l_1$, then a simple pendulum with length l_1 will have the same period as the actual physical pendulum. The point G has long been known as the center of oscillation. We note that since $k \neq 0$, G lies on the opposite side of C from O. Moreover, suppose we were to suspend the body by the point G; the new moment of inertia about the axis of suspension would be

$$I' = mk^{2} + m(l_{1} - l)^{2}$$

= $mk^{2} + ml_{1}^{2} + ml^{2} - 2ml_{1}l$.

But since $ll_1 = k^2 + l^2$, the above reduces to

Ì

$$I' = ml_1^2 - m(k^2 + l^2)$$

$$= \frac{mk^2}{l^2} (k^2 + l^2), \qquad (7.5-11)$$

so that the frequency of the resulting motion is

$$\nu_1 = \frac{1}{2\pi} \sqrt{\frac{mg(l_1 - l)}{mk^2}} = \frac{1}{2\pi} \sqrt{\frac{gl}{k^2 + l^2}}, \quad (7.5-12)$$

which, however, is precisely equal to ν from eq. (7.5–5); in other words the frequency is the same whether the body is suspended from axes passing through O or G, the center of suspension or center of oscillation. These points are thus interchangeable. It is not difficult to show that a blow transverse to the line OG at G transfers no momentum to the axis of suspension at O. Hence G is also called the center of percussion.

To revert to eq. (7.5-10), if the frequency is given, the length of the equivalent simple pendulum is fixed and hence (7.5-10) is

a quadratic equation in l. Solving we have

$$l = \frac{l_1 \pm \sqrt{l_1^2 - 4k^2}}{2}; (7.5-13)$$

there are then two values of l corresponding to each frequency, viz., $l' = \frac{l_1 + \sqrt{l_1^2 - 4k^2}}{2}$ and $l'' = \frac{l_1 - \sqrt{l_1^2 - 4k^2}}{2}$. Consequently

for any body there are two possible distances from the center of mass such that suspension at points at these distances yields the same frequency. It is a fact of considerable importance that the period and frequency of the simple pendulum are independent of the mass of the pendulum bob and of the substance of which the latter is made. This recalls a historical matter of some interest. namely the experiments of Newton using the pendulum to show that gravity gives the same acceleration to all rigid bodies independently of shape, constitution or mass. We must remember. however, that in any actual experiment the observed frequency will be influenced by a number of factors not accounted for in the simple theory presented in the previous section. Thus the air will provide a certain resistance to the motion (cf. Sec. 2.6). over, changes in temperature will affect l, and the finite size of the amplitude necessitates corrections to the simple assumption involved in $\sin \theta = \theta$ (see Sec. 8.1).

One of the most interesting physical uses of the pendulum is the evaluation of g, the acceleration of gravity. Consulting eq. (7.5-10) and combining it with (7.5-9) we have

$$k^2 + l^2 = \frac{lP^2g}{4\pi^2}$$
,

whence

$$g = \frac{4\pi^2(k^2 + l^2)}{P^2l}$$
, (7.5–14)

in terms of experimentally observable quantities. Of course the implication is that k^2 can be calculated from a geometrical knowledge of the body concerned. Since this is often difficult to carry out exactly, it is perhaps better to use two different axes of suspension, corresponding to the two values of l, viz., l_A and l_B , with corresponding periods, P_A and P_B respectively. Then in place

of (7.5-14) we have the two equations

$$k^2 + l_A{}^2 = \frac{l_A g P_A{}^2}{4\pi^2},$$
 (7.5–15)

$$k^2 + l_B{}^2 = \frac{l_B g P_B{}^2}{4\pi^2},$$
 (7.5–16)

whence elimination of k gives finally

$$g = \frac{4\pi^2[l_A^2 - l_B^2]}{l_A P_A^2 - l_B P_B^2}$$
 (7.5–17)

A pendulum used in the above fashion is called a *reversible* pendulum. It was used by Kater in a careful determination of g.

7.6. Plane Motion of a Rigid Body. Next to rotation about a fixed axis the most simple case of motion of a rigid body is that in which all its particles move in planes parallel to a fixed plane. This will in general involve both translation and rotation. As a special case consider the rolling under gravity of a homogeneous

right circular cylinder down a perfectly rough plane, on which no slipping can take place. The situation is schematically depicted in Fig. 7·10, which shows the cylinder of mass m and radius a in contact with the plane at A. The inclination of the plane to the horizontal is θ . The forces acting on the cylinder are the weight mg acting at the center of mass O, the reaction R of the plane at A, and the force of friction F, acting up

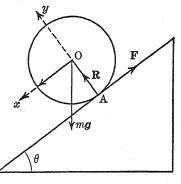


Fig. 7.10

the plane if the cylinder rolls down. When the cylinder rolls the center of mass O translates, while the particles composing the cylinder rotate about O. The motion therefore is a combination of translation and rotation which, as we have already seen (Sec. 7·1), is true in general for rigid bodies. We are therefore led to suppose that we can legitimately treat the two motions separately (cf. Secs. 6·3 and 7·7).

If we denote the displacement of the center of mass parallel to the plane by x, its equation of motion becomes

$$m\ddot{x} = mg\sin\theta - F. \tag{7.6-1}$$

Now the resultant torque about O is due wholly to \mathbf{F} , and we have therefore for the rotational motion about O,

$$Fa = I\alpha, (7.6-2)$$

where I is the moment of inertia about the axis of the cylinder, and α is the angular acceleration about this axis.

Next we note the purely kinematic relation

$$\ddot{x} = a\alpha. \tag{7.6-3}$$

The combination of (7.6-1), (7.6-2) and (7.6-3) yields at once

$$\ddot{x} = \frac{mg\sin\theta}{m + I/a^2} \cdot \tag{7.6-4}$$

But from Sec. 7.4 we know that the moment of inertia of a homogeneous right circular cylinder about its axis is $ma^2/2$. Hence the acceleration of the center of mass becomes

$$\ddot{x} = \frac{2}{3}g\sin\theta,\tag{7.6-5}$$

whose relation with the acceleration of a sliding object of an inclined plane should be carefully noted, as well as the fact that the dimensions or constitution of the cylinder nowhere appear in the formula.

It is interesting to observe that the problem of the rolling cylinder can also be solved by energy considerations. Conservation of energy dictates that the gain in kinetic energy of the rolling cylinder shall be equal to the loss in potential energy. Here we must, however, remember that the kinetic energy consists of two parts, i.e., kinetic energy of translation of the center of mass and kinetic energy of rotation about the center of mass. If, for simplicity, we suppose the cylinder starts from rest and rolls a distance x through a vertical drop $h=x\sin\theta$, the energy equation takes the form

$$\frac{1}{2}mv^2 + \frac{1}{2}I\omega^2 = mgx\sin\theta. {(7.6-6)}$$

Once again we have the kinematic relation between v and ω , viz.,

$$v = \omega a. \tag{7.6-7}$$

Hence

$$v^2 = \frac{2mgx \sin \theta}{m + I/a^2} = 2x \cdot \frac{2}{3}g \sin \theta.$$
 (7.6-8)

If we recall Sec. 2·1, it is clear that the constant acceleration of the center of mass is once more

$$\ddot{x} = \frac{2}{3}g\sin\theta.$$

7.7. General Equations of Motion of a Rigid Body. Having so far studied the nature of the motion of a rigid body and examined some special illustrations of a simple character, we now find it advisable to set up the equations of motion in general form. Actually these are included in the general equations for a collection of particles as treated in Chapter VI, but we shall repeat as much of the former analysis as seems necessary to give a complete account.

We shall agree that the translation of the rigid body is given by the motion of some point in it and take this as our origin. We are therefore here principally interested in the rotation of the body about this point. As we have seen in Sec. 6.3, the instantaneous moment of momentum about this origin of the jth particle in the body is given by

$$\mathbf{r}_{j} \times m_{j} \dot{\mathbf{r}}_{j}$$
.

But since the motion is that of a rigid body about the origin, we can at once replace $\dot{\mathbf{r}}_i$ by $\omega \times \mathbf{r}_i$ from (7.2–3) and write for the total moment of momentum of the rigid body about the origin the vector

$$\mathbf{M} = \sum_{j=1}^{n} m_{j} \mathbf{r}_{j} \times (\boldsymbol{\omega} \times \mathbf{r}_{j}), \qquad (7.7-1)$$

it being supposed that there are n particles in the body. The fundamental equation (6.4–10) still applies and hence we may write

$$\dot{\mathbf{M}} = \mathbf{L}, \tag{7.7-2}$$

which says that the time rate of change of the total moment of momentum, now to be renamed the angular momentum of the

rigid body, about the chosen origin is equal to the resultant torque of all the forces acting on the body about the chosen origin. This is the fundamental equation of rotational motion of the rigid body.

Let us look a bit more carefully at M. From $(7\cdot2-4)$ and the fact that $\mathbf{r}_j = \mathbf{i}x_j + \mathbf{j}y_j + \mathbf{k}z_j$, we can expand $\omega \times \mathbf{r}_j$ and get (cf. 6·4-4)

$$\omega \times \mathbf{r}_{j} = (\mathbf{i}\omega_{x} + \mathbf{j}\omega_{y} + \mathbf{k}\omega_{z}) \times (\mathbf{i}x_{j} + \mathbf{j}y_{j} + \mathbf{k}z_{j})$$
$$= \mathbf{i}(\omega_{y}z_{j} - \omega_{z}y_{j}) + \mathbf{j}(\omega_{z}x_{j} - \omega_{x}z_{j}) + \mathbf{k}(\omega_{x}y_{j} - \omega_{y}x_{j}).$$

Therefore

$$\mathbf{r}_{j} \times (\boldsymbol{\omega} \times \mathbf{r}_{j}) = \mathbf{i}[y, \ \omega_{x}y_{j} - \omega_{y}x_{j}) - z_{j}(\omega_{x}x_{j} - \omega_{x}z_{j})]$$

$$+ \mathbf{j}[z_{j}(\omega_{y}z_{j} - \omega_{x}y_{j}) - x_{j}(\omega_{x}y_{j} - \omega_{y}x_{j})]$$

$$+ \mathbf{k}[x_{j}(\omega_{x}x_{j} - \omega_{x}z_{j}) - y_{j}(\omega_{y}z_{j} - \omega_{x}y_{j})]. (7.7-3)$$

This leads to the following expanded expression for the angular momentum

$$\mathbf{M} = \mathbf{i}[\omega_x \sum m_j (y_j^2 + z_j^2) - \omega_y \sum m_j x_j y_j - \omega_z \sum m_j x_j z_j]$$

$$+ \mathbf{j}[-\omega_x \sum m_j x_j y_j + \omega_y \sum m_j (x_j^2 + z_j^2) - \omega_z \sum m_j y_j z_j]$$

$$+ \mathbf{k}[-\omega_x \sum m_j x_j z_j - \omega_y \sum m_j y_j z_j + \omega_z \sum m_j (x_j^2 + y_j^2)]. \quad (7.7-4)$$

By the introduction of simplifying notation we may write this in the form

$$\mathbf{M} = \mathbf{i}(\omega_x I_{xx} - \omega_y I_{xy} - \omega_z I_{xz})$$

$$+ \mathbf{j}(-\omega_x I_{yx} + \omega_y I_{yy} - \omega_z I_{yz})$$

$$+ \mathbf{k}(-\omega_x I_{zx} - \omega_y I_{zy} + \omega_z I_{zz}), \qquad (7.7-5)$$

where evidently I_{xx} , I_{yy} , I_{zz} from their makeup are the moments of inertia of the rigid body with respect to the x, y, z axes respectively. The quantities I_{xy} , I_{yz} , etc., are less familiar and we have not had to mention them in our earlier analysis. I_{xy} is termed the product of inertia with respect to the x and y axes, etc. Clearly we have $I_{xy} = I_{yx}$, etc.

Eq. (7.7-2) reduces to component equations of which the one corresponding to the x axis is

$$\dot{\omega}_x I_{xx} - \dot{\omega}_y I_{xy} - \dot{\omega}_z I_{xz} + \omega_x \dot{I}_{xx} - \omega_y \dot{I}_{xy} - \omega_z \dot{I}_{xz} = L_x, \quad (7.7-6)$$

where L_x is the component of the resultant torque about the x axis.

Note that the moments and products of inertia being calculated with respect to axes fixed in space will in general change with the time. Hence the general component equations of rotational motion are complicated. Obviously if the motion of the rigid body is restricted to rotation about the fixed x axis so that $\omega_y = \omega_z = 0$ at all times, eq. 7.7–6 reduces to

$$\frac{d}{dt}\left(I_{xx}\omega_{x}\right)=L_{x},$$

i.e., of the same form as (7.3-6).

We have already shown in Secs. 6·1 and 6·4 that the motion of a general collection of particles subject to arbitrary external and internal forces can be reduced to the motion of the center of mass, given by eq. (6·4–13) and the motion relative to the center of mass, given by eq. (6·4–14). This result holds, of course, for a rigid body as a special collection of particles. In the case of a rigid body, however, motion relative to the center of mass means rotational motion about the center of mass. Hence we can now state without further demonstration that the motion of a rigid body can be completely described by: (1), the translational motion of the center of mass, given by the equation

$$m\ddot{\ddot{\mathbf{r}}} = \mathbf{F}, \qquad (7.7-7)$$

where F denotes the vector sum of all the forces applied to the body, m is the mass of the body, and \bar{r} the position vector of the center of mass, and (2), the rotational motion about the center of mass, given by

$$\dot{\mathbf{M}}_c = \mathbf{L}_c, \qquad (7.7-8)$$

where \mathbf{M}_c is the angular momentum measured with respect to the center of mass as origin, and \mathbf{L}_c is the resultant torque due to the applied forces about the center of mass. In other words we can apply (7.7-8) to the center of mass just as if it were *fixed*. This may be called the principle of the independence of translational and rotational motions of a rigid body already illustrated in the treatment of plane motion in Sec. 7.6. It focusses attention on the great importance of the center of mass of a rigid body and hence makes the determination of the latter point of considerable significance. Before we go on to discuss examples of eq. (7.7-8) it will therefore be well to give some attention to the center of mass.

7.8. Center of Mass of a Rigid Body. The center of mass of a discrete collection of particles has already been defined as the point with the position vector [eq. (6.1-7)]

$$\bar{\mathbf{r}} = \frac{\sum m_i \mathbf{r}_i}{m},\tag{7.8-1}$$

where \mathbf{r}_i is the position vector of the *i*th particle, and m_i is its mass. The mass of the whole collection is $m = \sum m_i$. We have discussed an illustration of $(7\cdot8-1)$ in the two-particle problem of Sec. 6·3. Most rigid bodies, however, have to be considered as essentially continuous collections of particles and therefore $(7\cdot8-1)$ must be generalized as far as its mathematical calculation is concerned, though its physical meaning remains the same. From integral calculus we know that we can replace the limit of a sum such as that in $(7\cdot8-1)$ as the number of particles grows very large and the mass of each grows very small by a definite integral taken over the whole body. Thus, to go over to the rectangular coordinates, we must now replace $(7\cdot8-1)$ by

$$\overline{x} = \int x \, dm / \int dm,$$

$$\overline{y} = \int y \, dm / \int dm,$$

$$\overline{z} = \int z \, dm / \int dm.$$
(7.8-2)

Here dm denotes the mass of the element of volume of the body whose rectangular coördinates are x, y, z. Though no limits are indicated, it is understood that the integration is to be conducted over the whole volume of the body. We denote the mass m by $\int dm$.

A simple illustration of (7.8-2) is provided by a homogeneous infinitely thin rod of line density ρ , extending along the x axis a distance l from the origin. We at once have $\bar{y} = \bar{z} = 0$, whereas x is calculated by taking $dm = \rho dx$, where dx is the element of length at distance x from the origin. We then have

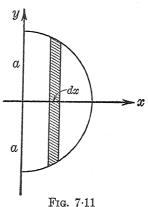
$$\bar{x} = \int_0^l \rho x \, dx / \int_0^l \rho \, dx = l/2.$$
 (7-8-3)

The center of mass is at the geometrical center of the rod, which is indeed intuitively clear from symmetry considerations.

We are often guided by symmetry in determining the center of

mass of a rigid body without evaluating integrals. A simple example is a homogeneous sphere for which the center of mass is evidently at the center. We can make the general statement that the center of mass of any homogeneous rigid body lies at its geometrical center or center of symmetry.

As another example of the general formulas (7.8-2) let us calculate the center of mass of a semicircular, homogeneous flat plate of radius a. We take the axes as in Fig. 7·11. From



. . .

symmetry the center of mass must lie on the x axis, i.e., as we have chosen our axes

$$\overline{y} = 0,$$

$$\overline{x} = \frac{\int x \, dm}{\int dm}$$
(7.8-4)

Now $dm = \rho dA$, where dA is an area element and ρ is the superficial density which is constant here. Then \bar{x} reduces to the center of area (or *centroid*, as it is often called), viz.,

$$\bar{x} = \frac{\int x \, dA}{A} \,, \tag{7.8-5}$$

where A is the area of the plate. The problem now is essentially mathematical in nature: the appropriate choice of the area element dA and the carrying out of the integration. For example, if we choose as our element the strip of width dx as illustrated in the figure,

$$dA = 2\sqrt{a^2 - x^2} dx.$$

Therefore

$$\bar{x} = \frac{2\int_0^a x \sqrt{a^2 - x^2} \, dx}{\pi a^2 / 2} = \frac{4a}{3\pi} \cdot \tag{7.8-6}$$

For z we have, of course, by symmetry h/2, where h is the thickness of the plate.

As a second illustration, let us consider the center of mass of a semicircular infinitely thin wire of uniform line density and radius a. With axes as before the equation of the circle with center at the origin is

$$x^2 + y^2 = a^2. (7.8-7)$$

In this case symmetry again tells us that the center of mass is on the line y=0 (the semicircular arc being the periphery of the semicircle considered previously). We have $dm=\rho ds$ where ρ is the line density (i.e., mass per unit length) and ds is the element of arc of the circle. Now

$$ds = \sqrt{1 + \left(\frac{dy}{dx}\right)^2} \, dx,\tag{7.8-8}$$

and from eq. (7.8-7)

$$\frac{dy}{dx} = -\frac{x}{y}.$$

On substitution into the expression for \bar{x} , we finally obtain

$$\bar{x} = \frac{2}{\pi} \int_0^a \frac{x \, dx}{\sqrt{a^2 - x^2}}$$

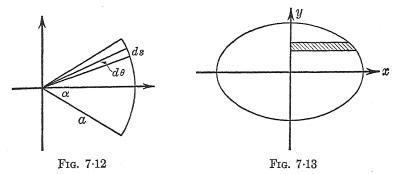
$$= \frac{2a}{\pi} \cdot \tag{7.8-9}$$

The same result can be achieved more simply by the use of polar coördinates. Let us, for example, take the more general case of any circular arc subtending the angle 2α at the center (see Fig. 7·12). Then $ds = a d\theta$ and $x = a \cos \theta$. Hence

$$\bar{x} = \frac{\int_{-\alpha}^{+\alpha} a \cos \theta \cdot a \, d\theta}{\int_{-\alpha}^{+\alpha} a \, d\theta} = a \frac{\sin \alpha}{\alpha} \cdot \tag{7.8-10}$$

For $\alpha = \pi/2$, this yields $\bar{x} = 2a/\pi$, the special result (7.8–9).

Cases where the body considered is non-homogeneous (i.e., of variable density) are sometimes important. For example, let us find the center of mass of the quadrant of an elliptical plate of constant thickness enclosed by the two semi-axes (Fig. 7·13). The



density is supposed to vary in such a way that at any point it is directly proportional to the distance from the point to the major axis. Symbolically, $\rho = ky$. Hence

$$\bar{x} = \frac{\int \rho x \, dA}{\int \rho \, dA} = \frac{\int yx \, dA}{\int y \, dA}, \qquad (7.8-11)$$

while

$$\bar{y} = \frac{\int y^2 dA}{\int y dA}$$
 (7.8–12)

To get \bar{y} , choose $dA = x dy = a/b \cdot \sqrt{b^2 - y^2} dy$ (the equation of the ellipse being $x^2/a^2 + y^2/b^2 = 1$). Then

$$\bar{y} = \frac{\frac{a}{b} \int_0^b y^2 \sqrt{b^2 - y^2} \, dy}{\frac{a}{b} \int_0^b y \sqrt{b^2 - y^2} \, dy} = \frac{3\pi b}{16} \cdot \tag{7.8-13}$$

To get \bar{x} , let us note that the center of mass of each of the horizontal strips, one of which is indicated in the figure, is x/2 which

equals $a/2b \cdot \sqrt{b^2 - y^2}$. Thus the whole mass of each strip $(ky \cdot dy \cdot a/b \cdot \sqrt{b^2 - y^2})$ may be considered to be located at the distance $a/2b \cdot \sqrt{b^2 - y^2}$ from the y axis. Therefore

$$\bar{x} = \frac{\frac{1}{2} \int_0^b \frac{a}{b} \cdot \sqrt{b^2 - y^2} \cdot \frac{a}{b} \cdot \sqrt{b^2 - y^2} \cdot y \, dy}{\frac{a}{b} \cdot \frac{1}{3} \cdot b^3}$$

$$= \frac{3}{8} a. \tag{7.8-14}$$

The center of mass of the quadrant is then the point $(\frac{3}{8}a, \frac{3\pi}{16}b)$.

So far we have confined our attention to the center of mass of plane plates or lamina. It remains to consider surfaces in general and volumes. The center of mass of any surface will be given by

$$\bar{x} = \frac{\int \sigma x \, dS}{\int \sigma \, dS}, \quad \bar{y} = \frac{\int \sigma y \, dS}{\int \sigma \, dS}, \quad \bar{z} = \frac{\int \sigma z \, dS}{\int \sigma \, dS}, \quad (7.8-15)$$

where dS is the element of area and σ is the mass per unit area (surface density). Perhaps the most important case of this kind is that of a surface of revolution. Let the x axis be the axis of symmetry. Then $\bar{y} = \bar{z} = 0$, and we have for the area element

$$dS = 2\pi y \sqrt{1 + \left(\frac{dy}{dx}\right)^2} dx, \qquad (7.8-16)$$

where y = f(x) is the equation of the generating curve. We can use this formula to obtain, for example, the center of mass of the curved surface of a right circular cone. Letting the vertex be at the origin we have

$$y = f(x) = bx,$$
 (7.8–17)

where b is the slope of the generator. For constant surface density there results

$$\bar{x} = \frac{\int 2\pi yx \sqrt{1 + \left(\frac{dy}{dx}\right)^2} dx}{\int 2\pi y \sqrt{1 + \left(\frac{dy}{dx}\right)^2} dx},$$
 (7.8–18)

from which on substitution we obtain

$$\bar{x} = \frac{\int_0^h x^2 \sqrt{1+b^2} \, dx}{\int_0^h x \sqrt{1+b^2} \, dx} = \frac{2}{3} h, \tag{7.8-19}$$

where h is the height of the cone.

For the volume enclosed by a surface we have

$$\overline{x} = \frac{\int \rho x \, dV}{\int \rho \, dV}, \quad \text{etc.}, \tag{7.8-20}$$

where ρ is the volume density and dV the volume element. Thus for the special case of the volume of the right circular cone above considered,

$$dV = \pi y^2 dx = \pi b^2 x^2 dx, \qquad (7.8-21)$$

and

$$\bar{x} = \frac{\int_0^h \pi b^2 x^3 \, dx}{\int_0^h \pi b^2 x^2 \, dx} = \frac{3}{4} h. \tag{7.8-22}$$

The difference between (7.8-22) and (7.8-19) should be noted.

The more general problem of finding the center of mass for any solid introduces the general volume element. In rectangular coördinates

$$dV = dx dy dz, (7.8-23)$$

and the calculation involves a triple integration. As a type problem let us consider the center of mass of that portion of the ellipsoid with equation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1, (7.8-24)$$

which is included in one octant, as is indicated in Fig. 7.14. We must choose the limits of integration of x, y, z. At x draw a thin slice dx parallel to the yz plane. Its volume is dx times one-

quarter of the area of the ellipse cut out by a plane parallel to the yz plane at this point. Since the area of an ellipse is π times the

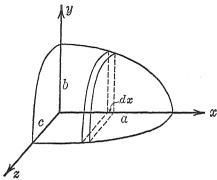


Fig. 7.14

product of the semi-major and semi-minor axes, the volume in question is $\frac{\pi}{4}$ yz dx, where, however,

$$y = b\sqrt{1 - \frac{x^2}{a^2}}, \quad z = c\sqrt{1 - \frac{x^2}{a^2}}.$$
 (7.8–25)

Therefore

$$\bar{x} = \frac{\int_0^a \frac{\pi}{4} bcx \left(1 - \frac{x^2}{a^2}\right) dx}{\int_0^a \frac{\pi bc}{4} \left(1 - \frac{x^2}{a^2}\right) dx}$$
$$= \frac{3}{8} a. \tag{7.8-26}$$

Similarly

$$\bar{y} = \frac{3}{8}b, \quad \bar{z} = \frac{3}{8}c.$$
 (7.8–27)

The reader should carry out the problem for the case where the density is variable.

7.9. Equilibrium of a Rigid Body. Center of Gravity. In Chapter V we agreed to say that a particle is in equilibrium under the action of a set of forces if its acceleration with respect to the

primary inertial system vanishes. We have now to decide what we shall mean by the equilibrium of a rigid body. Clearly if the center of mass of the body remains at rest or moves with constant velocity in a straight line it will be appropriate to say that the body is in translational equilibrium. In this condition, however, it still might suffer rotational acceleration about the center of mass and in this case we should not wish to say it is in equilibrium with respect to rotation. The natural definition of the latter is the absence of rotational acceleration about any axis through the center of mass or indeed any other point; this will ensue (from

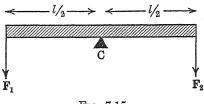


Fig. 7.15

eq. (7.7-8)) if the resultant torque of the applied forces about the center of mass vanishes. In this case it may readily be shown that the torque will also vanish about any axis whatever, provided the forces also satisfy the condition of translational equilibrium. There are thus two conditions of equilibrium of a rigid body.

A simple case of a rigid body in translational equilibrium but not necessarily in rotational equilibrium is provided by a homogeneous rod of length l supported by a fulcrum at its center of mass. This is shown schematically in Fig. 7·15. We imagine that forces \mathbf{F}_1 and \mathbf{F}_2 act in the same plane at the left and right ends of the rod, respectively. Clearly, as long as the fulcrum can exert an upward reaction force on the rod equal in magnitude to $|\mathbf{F}_1 + \mathbf{F}_2|$ the center of mass will remain at rest. However, the resultant torque about

¹ More careful consideration into which we shall not enter here (but cf. eq. 7·11–15 and accompanying discussion) discloses that this definition has defects. Actually it turns out that it is possible to have a rigid body rotate with angular acceleration about a certain axis if the resultant torque is zero, provided it already has an angular velocity to begin with. It is also possible to envisage a situation in which a resultant torque will accompany constant angular velocity about some axis. To avoid these difficulties we shall actually confine our association of the term rotational equilibrium with a rigid body to the case of rest.

the center of mass has the magnitude (in the clockwise direction)

$$F_2 l/2 - F_1 l/2$$
,

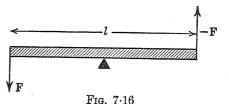
and this will in general produce a rotational acceleration about C (clockwise if $F_2 > F_1$ and counterclockwise if $F_2 < F_1$). To assure rotational equilibrium it is necessary to have

$$F_1 = F_2. (7.9-1)$$

Note that if F_2 is directed *upward* and F_1 directed *downward*, $(7\cdot9-1)$ could not produce equilibrium unless both forces have zero magnitude. The particular case

$$\mathbf{F}_1 = -\mathbf{F}_2 \tag{7.9-2}$$

is interesting. Here F_1 and F_2 are said to form a *couple*. A rigid body acted on by a couple will be in translational equilibrium so



far as the forces of the couple are concerned but cannot be in rotational equilibrium.

From Fig. 7-16 it is clear that the torque produced by the couple about the center of mass C (strictly the torque about an axis through C perpendicular to the plane of the forces) has the magnitude

Fl.

This is often termed the *moment* of the couple. The perpendicular distance (in this case l) between the lines of action of the forces of the couple is termed the arm of the couple. The reader may show that the moment of the couple has the same value about any axis perpendicular to the plane. It is also simple to prove that, so far as its rotational action on a rigid body is concerned, a couple may be replaced by any other couple with the same moment in the same plane.

Since a couple involves parallel forces in a plane, it will be well to consider the general problem of the composition of coplanar parallel forces acting on a rigid body. Consulting Fig. 7·17 let us assume that the coplanar forces F_1 , F_2 and F_3 act at points O_1 , O_2 and O_3 , respectively, of the body schematically indicated. Now the resultant of the forces is the vector $F_1 + F_2 + F_3$ with magnitude $F_1 + F_2 + F_3$. Unfortunately the method of vector summation does not prescribe the line of action of the resultant and hence in this case we must settle this by arbitrary definition. Here it is appropriate to agree that the resultant force must produce the

same translational and rotational effects as the forces individually. The former effect is, of course, independent of the resultant line of action. The latter, however, demands that the moment or torque of the resultant about any axis perpendicular to the plane shall be equal to the sum of the moments of the individual forces

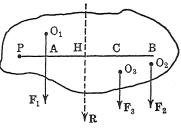


Fig. 7.17

about the same axis. Consider any point P and draw the line PACB perpendicular to the three forces and intersecting their lines of action at A, B, C, respectively, with PA = a, PB = b and PC = c. We now must suppose the resultant \mathbf{R} to be drawn parallel to the individual forces and at distance PH = h from P such that

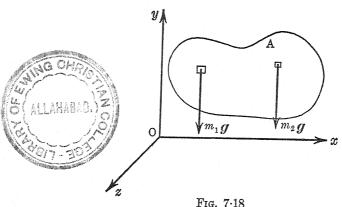
$$h = \frac{aF_1 + bF_2 + cF_3}{F_1 + F_2 + F_3}$$
 (7.9-3)

This fixes the line of action of R, although it does not determine its point of application.

The most important case of parallel, but not necessarily coplanar, forces acting on a rigid body is provided by the force of gravity. In Fig. 7-18 we imagine a rigid body referred to the set of axes indicated. We think of it as composed of the collection of mass particles m_1, m_2, m_3, \ldots possessing weights m_1g, m_2g, m_3g, \ldots which are all parallel forces in so far as we can consider \mathbf{g} to have the same direction for all the particles of the body. If the body is very large, this is not strictly true, but is sufficiently so if the body is not too extended. Clearly,

$$\mathbf{W} = \sum m_i \mathbf{g} \tag{7.9-4}$$

is the weight of the body if the sum is extended over all its particles. Along what line now does the weight act? Let us denote the coördinates of the *i*th particle by x_i , y_i , z_i . If we apply the criterion



of the previous paragraph, it follows that the resultant W will lie in a plane parallel to the yz plane with perpendicular distance from the latter given by

$$\bar{x} = \sum m_i x_i g / mg, \qquad (7.9-5)$$

where

$$m = \sum m_i \tag{7.9-6}$$

is the total mass of the body. This results from the requirement that the moment of W about the z axis must be the sum of the separate moments about the z axis of all the individual particle weights. Taking moments in similar fashion about the x axis, we find that the resultant must also lie in a plane parallel to the xy plane and distant from the latter by

$$\bar{z} = \sum m_i z_i g / mg. \tag{7.9-7}$$

Finally if we turn the whole body around through an angle of 90° about the z axis, so that the gravity forces act parallel to the x axis, we find, taking moments once more about the z axis that the resultant must lie in a plane parallel to the zx plane and at distance from it

$$\bar{y} = \sum m_i y_i g / mg. \tag{7.9-8}$$

The three planes thus determined meet in the point $(\bar{x}, \bar{y}, \bar{z})$, which

is thus the center of the rigid body as far as the action of the parallel gravity forces is concerned. It is the point through which the resultant force of gravity always acts no matter how the body is placed and is termed therefore the center of gravity. The formulas (7.9-5), (7.9-7) and (7.9-8) suffice to determine its position with respect to any set of axes for a body made up of discrete parts. For a continuous body, however, the summations must be replaced by integrations. In this case we have for \bar{x} , \bar{y} , \bar{z} , precisely the equations (7.8-2) from which the g, being constant, has disappeared. What we have actually shown then is that the center of gravity for a body of extent small compared with the earth coincides approximately with the center of mass. For a rigid body of large extent the center of gravity can not be obtained by eqs. (7.8-2) since the weight forces on the various portions of the body will then no longer be parallel, since g is no longer to be considered constant either in magnitude or direction over the whole body. In this case the center of gravity and the center of mass do not The coincidence, however, will naturally be almost exactly true for any body used, for example, for engineering purposes on the earth's surface. Since the mass concept is independent of gravitation in classical mechanics and since the mass of any particle remains constant everywhere at all times, the center of mass of a rigid body is a more fundamental quantity than the center of gravity.

We may appropriately point out here that the center of gravity and center of mass are but two illustrations of the general concept of center of mean position with respect to any set of effects. For example, in a given country we can ascertain the center of *population* with coördinates x_p and y_p , where

$$x_p = \frac{\sum N_i x_i}{N}, \quad y_p = \frac{\sum N_i y_i}{N},$$

where N_i is the number of people in an arbitrarily chosen area element of the country with coördinates x_i and y_i with respect to some chosen origin, and N is the total number of people in the country. Other more physical illustrations are center of pressure (see Chap. XI), center of area and center of volume. In each case the problem is one of finding the *mean position* with respect to some property of the body or some outside influence on the body.

An interesting and important illustration of the concept of center of gravity, particularly with reference to parallel forces, is to be found in the common balance used for the measurement of mass. Consulting Fig. 7-19, imagine AB to represent schematically a rigid rod with center of gravity at C, a distance OC = a below the point O which is half-way between A and B, with AO = OB = l. Let the mass of the rod be m. Assume that the rod is supported at O and that at O and O and

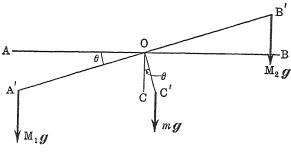


Fig. 7-19

respectively. Then the rod will assume for equilibrium a position A'B' making the angle θ with AB. If the rod is not too long the equilibrium may be considered as due to the parallel forces $M_1\mathbf{g}$, $m\mathbf{g}$ and $M_2\mathbf{g}$. (Strictly speaking these forces are never exactly parallel nor is the value of \mathbf{g} exactly the same for each mass. But since the rod is small compared with the earth, the approximation is an extraordinarily good one for all practical purposes.) Taking moments about O we have

$$g(M_1 - M_2)l\cos\theta - gma\sin\theta = 0$$

or

$$M_1 - M_2 = \frac{ma}{l} \tan \theta. \tag{7.9-9}$$

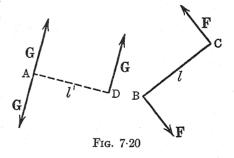
In this way the two masses M_1 and M_2 may be compared, and we have then a method of measuring mass which is of greater practical value than the ideal scheme on which the definition of mass used in this text has been based (Sec. 1-6). The reader will, however, easily convince himself that while the balance affords an accurate practical method of mass measurement, it does not provide a satisfactory method for defining mass.

There are several ways of using a balance for the attainment of maximum accuracy. The method of waiting for the attainment of the equilibrium indicated in eq. (7.9-9) is usually a slow process, for the balance will oscillate about the equilibrium position before coming to rest. A pointer is usually attached to the balance beam at O and is arranged to move over a fixed scale. Equality of the masses M_1 and M_2 is then assumed when the amplitudes of the pointer movements to the right and left of a point vertically under O are equal. This of course assumes that the arms (i.e., AO and OB) are really of the same length I. To avoid errors due to a possible difference here, Gauss suggested that the body whose mass is being measured be weighed first on one side and then on the other. If the two results are denoted by M' and M'', the reader may show that the actual mass is given by

$$M = \sqrt{M'M''}. (7.9-10)$$

7.10. Equilibrium under Coplanar Forces. Illustrations. The conditions of equilibrium of a rigid body can be most readily visualized when the forces acting on it lie in a plane. To proceed with this let us note an important result about couples: a co-

planar force and couple acting on a rigid body are equivalent in their action to a single force. Thus consider the couple consisting of the pair of forces F with the arm l, and the additional force G acting at A (cf. Fig. 7.20). Now the couple may be replaced by any other couple in the



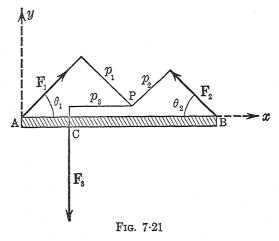
same plane with a moment equal in magnitude and direction without changing its effect on the body. Let us then replace the original couple by another with the extremities of its arm at A and D and with moment of magnitude

$$Gl' = Fl.$$

From the way the new but equivalent couple has been chosen, one of its forces is equal and opposite to G so that the two cancel each

other in their action on the rigid body, and there remains effectively the force G acting in a line parallel to the original force but at a perpendicular distance l' = Fl/G from the line of application of the latter. This proves the result stated above.

We can now appreciate the important fact that if a force and a couple act in a single plane they can not produce equilibrium. We can also use the above proposition to prove at once that a given force F acting at a point A of a rigid body can be replaced by an equal force acting at any other point B of the body, together



with a couple with moment equal to the moment of F about an axis through B perpendicular to the plane of the forces. Suppose now we have acting on a rigid body the coplanar forces F_1, F_2, \ldots, F_n . We proceed in the usual way to find a resultant (cf. Sec. 1·3), i.e., we translate the line of action of F_2 so that its origin is at the end of F_1 , and find the resultant of F_1 and F_2 which we may call R_{12} . We then compound this with F_3 to get R_{123} , etc., until we finally have $R_{123} \ldots n$, the sum of all the forces, viz.,

$$R_{12...n} = F_1 + F_2 + \cdots + F_n.$$
 (7·10-1)

However we must be careful to note the possibility that $R_{12...n-1}$ and F_n may form a *couple*, and hence can not be reduced to a single force whose effect on the rigid body is the same as that of all the n forces acting together. Consider for example the special case of a rigid rod AB of length l, and assume that the coplanar

211

forces $\mathbf{F_1}$, $\mathbf{F_2}$, and $\mathbf{F_3}$ act at A, B and C respectively, where, to be specific, $\theta_1 = \theta_2 = 45^\circ$ and AC = l/4; also $F_1 \sin \theta_1 + F_2 \sin \theta_2 = F_3$ and $F_1 \cos \theta_1 = F_2 \cos \theta_2$, i.e., the forces are in equilibrium with respect to translations (cf. Fig. 7.21). That is,

$$\sum \mathbf{F} = 0. \tag{7.10-2}$$

Let us take any point P in the plane of the rod and forces, and calculate the moment of the forces or torque about a perpendicular axis through this point keeping θ_1 and θ_2 perfectly general. Let the perpendicular distances from P to the lines of action of the forces be p_1 , p_2 and p_3 respectively. Then the resultant counterclockwise torque has the magnitude

$$-p_1F_1 + p_2F_2 + p_3F_3. (7.10-3)$$

Suppose the coördinates of P with respect to A as origin are x_1 , y_1 , taking the x axis along the rod. The equation of the line of action of \mathbf{F}_1 is $y = x \tan \theta_1$. (7·10-4)

That of the line of action of F2 is

$$y = -(x - l) \tan \theta_2,$$
 (7.10-5)

and finally that of the line of action of F3 is here equivalent to

$$x = \frac{l}{4} \cdot \tag{7.10-6}$$

Now the perpendicular distance from the point (x_1, y_1) to the line whose equation is

$$Ax + By + C = 0,$$
 (7·10–7)

is from analytic geometry given by

$$p = \frac{Ax_1 + By_1 + C}{\sqrt{A^2 + B^2}}$$
 (7·10-8)

On substitution we finally have in our special case

$$p_{1} = \frac{y_{1} - x_{1} \tan \theta_{1}}{\sec \theta_{1}},$$

$$p_{2} = \frac{x_{1} \tan \theta_{2} + y_{1} - l \tan \theta_{2}}{\sec \theta_{2}},$$

$$p_{3} = -x_{1} + \frac{l}{4}.$$
(7·10-9)

The expression for the resultant moment then becomes

$$-p_1F_1 + p_2F_2 + p_3F_3 = F_3\frac{l}{4} - F_2l\sin\theta_2, \quad (7.10-10)$$

i.e., a constant independent of the position of P. Moreover it is different from zero unless

$$F_3 = 4F_2 \sin \theta_2, \tag{7.10-11}$$

which in turn is incompatible with the assumed relation (7·10-2), unless we have the relation between the angles

$$3 \tan \theta_2 = \tan \theta_1. \tag{7.10-12}$$

The latter is the condition that, in addition to translational equilibrium, rotational equilibrium shall also be maintained. In general, however, as in the present case, this condition will not be satisfied and there will be a resultant moment so that the three forces will be equivalent to a couple with this moment. The arm of this couple may be chosen anywhere in the plane of the forces.

To summarize: If we have a system of forces in a plane they may either have a single force as a resultant or reduce to a *couple*.

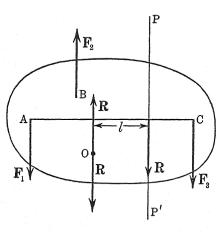


Fig. 7.22

Suppose now that the forces actually have a resultant. Consulting Fig. 7.22, where the three parallel forces \mathbf{F}_1 , F₂ and F₃ are represented as acting on a rigid body at the three points A, B and C respectively, we proceed to find the resultant R passing through the line PP'. Now it is sometimes convenient to consider this resultant R as acting at some other point, such as O in the figure. In order to do this we must, however, introduce at O another force

R equal in magnitude and opposite in direction. This will then not alter the situation. But the new force will form with the original force a couple of moment =Rl in magnitude, where l is

the perpendicular distance from O to PP'. Hence we have replaced the single resultant force by a force of equal magnitude acting at the point O and a couple of moment Rl, which in the example used is clockwise. By simple analysis it can be shown that the moment Rl is the sum of the moments of the original forces F_1 , F_2 and F_3 about the point O. Hence the system reduces to the force R and couple of moment M where

and
$$R = \sum F_{i},$$

$$M = \sum F_{i}l_{i},$$

$$(7 \cdot 10 - 13)$$

 l_i being the perpendicular distance from the line of action of the *i*th force to O. The reader should carry out the analysis proving the similar result for a system of coplanar, *non*-parallel forces.

Finally it can be shown¹ that any system of forces (not necessarily parallel or coplanar) acting on a rigid body can be replaced by a force acting at any particular chosen point and a couple.

It can further be proved that any system of forces acting on a rigid body can be replaced by a force and a couple whose axis is parallel to the line of action of the force.

It being now clear that any system of forces acting on a rigid body can be replaced by a single force acting at any arbitrarily chosen point and a couple, under what conditions will the body be in equilibrium? We have already noted that a single nonvanishing force and a couple in the same plane can not produce equilibrium. The same result follows even if the force and the couple are not coplanar. For we may always transfer the couple in its plane so that one of its forces, say F, intersects the line of action of the given force R. The resultant of the two forces F and R can not equilibrate the other force of the couple, and hence the general statement follows. It therefore results that for the rigid body to be in equilibrium the single force must be equal to zero and the moment of the couple must likewise vanish. If the former (i.e., the force) is denoted by R and the latter (i.e., the moment of the couple) by L, the conditions of equilibrium may be written in the form

$$R = 0,$$
 $(7.10-14)$

$$L = 0. (7.10-15)$$

¹ Cf. Jeans, "Theoretical Mechanics" (Ginn & Co., 1907), p. 106 ff.

As an illustration of the application of the principles of equilibrium to a rigid body, consider the simple case of a ladder AB resting with one end against a smooth wall, and the other on the

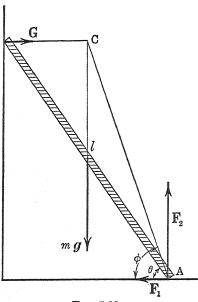


Fig. 7.23

ground. We suppose that the ladder is of uniform density. Hence by symmetry the center of mass is at the geometrical center. Its length is l, and it rests so that it makes an angle θ with the ground (see Fig. 7.23). The various forces acting on the ladder may then be tabulated as follows: (1) the reaction G of the wall, which, since the surface is smooth, is normal to the surface; (2) the reaction of the ground F which, on the other hand. will not be normal to the ground, for it is a rough reaction; (3) the weight W acting at the center of mass.

Since the direction of \mathbf{F} is unknown, we find it most simple to consider the horizontal and vertical components F_1 and F_2 as our two unknowns (though of course we could take \mathbf{F} and the angle ϕ it makes with the ground). We then proceed to write the conditions of equilibrium. First the condition for translational equilibrium, viz.,

$$\sum F_x = 0,$$

$$\sum F_y = 0.$$

These give respectively

$$W = mg = F_2, F_1 = G.$$
 (7·10-16)

The condition for rotational equilibrium we may write in a variety of ways, for the total moment of all the forces about any point in the plane is zero. We shall preferably choose the point in such a way as to render as small as possible the number of forces having a

215

moment different from zero. In the present case we shall thus naturally choose the point A.

$$Gl\sin\theta - mg\frac{l}{2}\cos\theta = 0,$$

or

$$G = \frac{mg}{2}\cot\theta = F_1. \tag{7.10-17}$$

The magnitude of F may be found at once.

$$F = \sqrt{m^2 g^2 + m^2 g^2 \frac{\cot^2 \theta}{4}} = mg \sqrt{1 + \frac{\cot^2 \theta}{4}}, \quad (7.10-18)$$

and the angle ϕ is given by

$$\phi = \arctan\left(\frac{2}{\cot\theta}\right) = \arctan\left(2\tan\theta\right).$$
 (7.10-19)

We may note an interesting thing about this problem. effectively three forces acting on the ladder, G, W and F. If we

extend the lines of action of these forces sufficiently they will meet in a point. This will be true of any three coplanar non-parallel forces in equilibrium. We can prove the theorem very simply. Let the lines of action of the forces G and F meet at the point C. Now let us take moments about an axis through C normal to the plane of the forces. The moments of G and F will of course be zero since the moment arms vanish. since the total moment of all the forces must be zero about C, that of W must be also. Hence the line of action of W must pass through C, and the theorem is proved. It is worth noting that we can use the latter to solve other problems involving the equilibrium of three non-parallel coplanar forces. we represent the angles made at the point

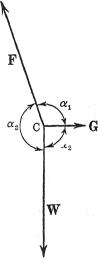


Fig. 7.24

C by the lines of action of the three forces with each other as α_1 , α_2 , α_3 respectively (Fig. 7.24), we have from the law of sines

$$\frac{W}{\sin \alpha_1} = \frac{F}{\sin \alpha_2} = \frac{G}{\sin \alpha_3}, \qquad (7.10-20)$$

as we have already noted in the previous chapter in the case of three forces acting at a point.

However, in most practical cases the reader will probably find the method used in eq. (7·10–16) (the so-called method of components) the most advantageous one for problems involving coplanar forces. It is of course not restricted to the case of three forces only, but always yields three independent equations from which three unknowns may be evaluated. The student should convince himself by trial that instead of the two equations $\sum F_x = 0$, $\sum F_y = 0$, we could use two other equations obtained by setting equal to zero the torque about two different axes (each different from the one chosen already for the eq. $\sum L = 0$).

7.11. Moving Axes. We now return to the problem of the motion of a rigid body, the general equations for which were set up in Sec. 7.7, and indeed expressed in most general form in eqs. (7.7-7) and (7.7-8). It will be recalled that these equations refer to axes fixed in space, i.e., the primary inertial system. We commented on the fact that as the body rotates about the center of mass (or any other point for that matter) the moments and products of inertia entering into the angular momentum M (eq. 7.7-5) will change, and that consequently the rotational equation of motion (7.7-8) will, in general, be very complicated. Much of this complexity can be removed by a rather simple expedient, namely, referring the rotational motion to axes which are fixed in the body and hence move with the body. We refer to these as moving axes and now wish to give some attention to them. respect to these axes the products and moments of inertia remain invariant, so that, in $\dot{\mathbf{M}}$, terms like \dot{I}_{xx} and \dot{I}_{xy} vanish. It must be pointed out, to be sure, that there is the usual compensation for this gain in simplification: we can now no longer look upon the unit vectors i, j, k as constant in time. As the body moves they will change, and this change will be reflected in M. Thus, denoting the components of M along the new axes (fixed in the body) as M_x , M_y , M_z , we must now write

$$\dot{\mathbf{M}} = \mathbf{i}\dot{M}_x + \mathbf{j}\dot{M}_y + \mathbf{k}\dot{M}_z + M_x\frac{d\mathbf{i}}{dt} + M_y\frac{d\mathbf{j}}{dt} + M_z\frac{d\mathbf{k}}{dt}. \quad (7.11-1)$$

Here the last three terms represent the contribution to $\dot{\mathbf{M}}$ due to the rotation of the axes. We proceed to evaluate $d\mathbf{i}/dt$, etc., on

the assumption that the origin of the moving axes is fixed in space, so that the only possible motion of the axes is one of rotation about a fixed point.

Since i is a unit vector, di/dt must be perpendicular to i and hence must lie in the plane of j and k. Therefore we can write

$$\frac{d\mathbf{i}}{dt} = A_3 \mathbf{j} - A_2 \mathbf{k}, \tag{7.11-2}$$

where A_3 and A_2 are coefficients which are initially undetermined. Similarly

$$\frac{d\mathbf{j}}{dt} = A_1\mathbf{k} - B_3\mathbf{i}; \quad \frac{d\mathbf{k}}{dt} = B_2\mathbf{i} - B_1\mathbf{j}. \tag{7.11-3}$$

But we recall from (6.4-4) that

$$i = j \times k$$
; $j = k \times i$; $k = i \times j$.

Hence

$$\frac{d\mathbf{i}}{dt} = \frac{d\mathbf{j}}{dt} \times \mathbf{k} + \mathbf{j} \times \frac{d\mathbf{k}}{dt}, \qquad (7.11-4)$$

and therefore from (7·11-2) and (7·11-3)

$$A_3 \mathbf{j} - A_2 \mathbf{k} = (A_1 \mathbf{k} - B_3 \mathbf{i}) \times \mathbf{k} + \mathbf{j} \times (B_2 \mathbf{i} - B_1 \mathbf{j}).$$
 (7.11-5)

From this it follows by comparing coefficients of identical unit vectors that

$$A_3 = B_3$$
 and $A_2 = B_2$. (7.11-6)

Similarly we can show that $A_1 = B_1$ so that the six coefficients are reduced to three independent ones. Hence finally

$$\frac{d\mathbf{i}}{dt} = A_3\mathbf{j} - A_2\mathbf{k}; \quad \frac{d\mathbf{j}}{dt} = A_1\mathbf{k} - A_3\mathbf{i}; \quad \frac{d\mathbf{k}}{dt} = A_2\mathbf{i} - A_1\mathbf{j}.$$
(7.11-7)

To see the significance of the coefficients A_1 , A_2 , A_3 , dot multiply $d\mathbf{i}/dt$ with \mathbf{j} etc., and get (recalling $\mathbf{k} \cdot \mathbf{j} = 0$, etc.)

$$\mathbf{j} \cdot \frac{d\mathbf{i}}{dt} = A_3; \quad \mathbf{k} \cdot \frac{d\mathbf{j}}{dt} = A_1; \quad \mathbf{i} \cdot \frac{d\mathbf{k}}{dt} = A_2.$$
 (7.11–8)

This means, for example, that A_3 is the component of $d\mathbf{i}/dt$ along the y axis and that $\mathbf{j} \cdot d\mathbf{i} = A_3 dt$ is the component of $d\mathbf{i}$ along the y axis. We represent this in Fig. 7-25 in which $\mathbf{j} \cdot d\mathbf{i}$ clearly is the

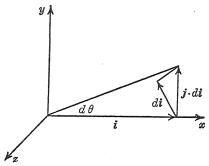


Fig. 7.25

angle $d\theta$ through which the x axis rotates about the z axis due to the change $d\mathbf{i}$ during the time dt. But $d\theta$ is $\omega_z dt$ by definition (cf. 7·2-4) and hence we have

$$A_3 = \omega_z. \tag{7.11-9}$$

Similarly it can be shown that

$$A_1 = \omega_x; \quad A_2 = \omega_y.$$
 (7.11-10)

We can now go back to (7.11-1) and write

$$\dot{\mathbf{M}} = i\dot{M}_x + j\dot{M}_y + k\dot{M}_z
+ i(M_z\omega_y - M_y\omega_z) + j(M_z\omega_z - M_z\omega_x)
+ k(M_y\omega_x - M_z\omega_y)
= i\dot{M}_x + j\dot{M}_y + k\dot{M}_z + \omega \times \mathbf{M},$$
(7.11-11)

utilizing the fact that

$$\mathbf{\omega} \times \mathbf{M} = (\mathbf{i}\omega_x + \mathbf{j}\omega_y + \mathbf{k}\omega_z) \times (\mathbf{i}M_x + \mathbf{j}M_y + \mathbf{k}M_z)$$

$$= \mathbf{i}(M_z\omega_y - M_y\omega_z) + \mathbf{j}(M_z\omega_z - M_z\omega_x)$$

$$+ \mathbf{k}(M_y\omega_x - M_z\omega_y) \quad (7.11-12)$$

by the use of the distributive rule and the relations (6.4-4).¹ It

¹ It should perhaps be emphasized that the method of expressing $\dot{\mathbf{M}}$ in terms of moving axes leading to (7·11–11) can be applied to any vector. This fact will be utilized in the following section.

is often valuable to express the cross product in determinant form to facilitate its expansion into components. Thus

$$\omega \times \mathbf{M} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \omega_x & \omega_y & \omega_z \\ M_x & M_y & M_z \end{vmatrix} . \tag{7.11-13}$$

Equation (7·11–11) is a general equation whose physical meaning will become clearer if we consider some special cases. First suppose that instead of a single point being fixed, there is a fixed line of points or axis and take this as the z axis. The x and y axes rotate then with the body about the z axis, and we have $\omega_z = \omega_y = 0$ and $\omega_z = \omega$. Consequently from (7·7–5)

$$\mathbf{M} = \omega(-iI_{xz} - jI_{yz} + kI_{zz}),$$
 (7.11-14)

and the equation of motion (7.7-2) (or (7.7-8) if the fixed axis passes through the center of mass) becomes¹

$$\mathbf{L} = \mathbf{i}(-I_{xz}\dot{\omega} + I_{yz}\omega^2) + \mathbf{j}(-I_{yz}\dot{\omega} - I_{xz}\omega^2) + \mathbf{k}I_{zz}\dot{\omega}. \quad (7.11-15)$$

If we specialize still further and suppose that the body is symmetrical about a plane perpendicular to the axis of rotation, $I_{yz} = I_{zz} = 0$ and (7.11-15) reduces to

$$\mathbf{L} = \mathbf{k} I_{zz} \dot{\omega}, \tag{7.11-16}$$

which is recognized as equivalent to the equation (7.3-7) for rotational acceleration about a fixed axis.

7.12. More About Moving Axes. Motion of a Particle on the Earth's Surface. The use of moving axes is of sufficient importance to warrant further consideration from a somewhat more general point of view, which is indeed not restricted to rigid bodies but may refer equally well to particles in general. In Fig. 7.26 we represent a set of rectangular axes, fixed with respect to the primary inertial system, in which the position of the point P is given by the coördinates x_f , y_f , z_f , with origin at O_f . Similarly we introduce a set of moving rectangular axes x_m , y_m , z_m with center at O_m , whose position vector in the fixed system is \mathbf{r}_0 . This,

¹ We note from (7·11–15) the interesting fact that, if the body has at any instant an angular velocity ω, there will also exist an angular acceleration ω even if L=0. This has a bearing on the difficulty of defining the rotational equilibrium of a rigid body already mentioned in Sec. 7·9. Of course, the difficulty vanishes if the rotation takes place about an axis of symmetry.

of course, can change with the time. The position vector of a particle at P in the moving system is \mathbf{r}_m while in the fixed system it is \mathbf{r}_f . From the figure we have

$$\mathbf{r}_m = \mathbf{r}_f - \mathbf{r}_0,$$
 (7.12-1)

and likewise for the corresponding velocities

$$\dot{\mathbf{r}}_m = \dot{\mathbf{r}}_f - \dot{\mathbf{r}}_0. \tag{7.12-2}$$

Let the unit vectors in the fixed system be i_f , j_f , k_f and those in the moving system be i_m , j_m , k_m . The former remain unchanged in

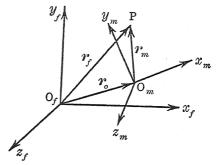


Fig. 7.26

time, whereas the latter change with the time, since the directions of the moving axes vary with time with respect to the fixed axes. Hence we can write in analogy with (7.11-1)

$$\dot{\mathbf{r}}_{m} = \mathbf{i}_{m}\dot{x}_{m} + \mathbf{j}_{m}\dot{y}_{m} + \mathbf{k}_{m}\dot{z}_{m} + x_{m}\frac{d\mathbf{i}_{m}}{dt} + y_{m}\frac{d\mathbf{j}_{m}}{dt} + z_{m}\frac{d\mathbf{k}_{m}}{dt}.$$
(7.12–3)

The first three terms yield the apparent velocity of P relative to the moving axes, whereas the last three terms represent the contribution to the velocity of P due to the rotation of the moving axes. We have already worked out the expressions for $d\mathbf{i}_m/dt$, etc., in Sec. 7·11 and hence can replace $(7\cdot12-3)$ by

$$\dot{\mathbf{r}}_m = \dot{\mathbf{r}}_{ma} + \mathbf{\omega} \times \mathbf{r}_m, \tag{7.12-4}$$

where $\dot{\mathbf{r}}_{ma}$ is the apparent velocity of P in the moving system. We need, however, to make sure of the meaning of ω . Actually it

is the angular velocity of the moving axes about the instantaneous axis of rotation through O_m . We shall show first that ω has the direction of the instantaneous axis of rotation through O_m . Since this axis is the locus of points which have at the instant in question zero linear velocity as far as rotation is concerned, if P lies on the axis through O_m so that \mathbf{r}_m is the axis, we must have from (7.12-3)

$$x_m \frac{d\mathbf{i}_m}{dt} + y_m \frac{d\mathbf{j}_m}{dt} + z_m \frac{d\mathbf{k}_m}{dt} = 0.$$
 (7·12–5)

But this means

$$\omega \times \mathbf{r}_m = 0, \qquad (7.12-6)$$

whence ω and \mathbf{r}_m are parallel. Hence ω has the same direction as the axis of rotation through O_m . We can next show that the magnitude of ω is equal to that of the angular velocity of the axes about the instantaneous axis through O_m . Consider a point P not on the axis of rotation. Its translational velocity due to the rotation of the axes has the magnitude $\omega r_m \sin \theta$, where θ is the angle between \mathbf{r}_m and ω . Moreover, the direction of the translational velocity is at right angles to the plane of ω and \mathbf{r}_m . This velocity is therefore precisely that of a point moving in a circle of radius $r_m \sin \theta$ with angular speed ω about the center, which in this case lies on the axis of rotation. Finally, the sign of ω agrees with that of the rotation of the axes as is clear from $(7\cdot12-4)$.

We can now go back to $(7\cdot12-2)$ and write for the general relation between velocity in the fixed system and that in the moving system

$$\dot{\mathbf{r}}_f = \dot{\mathbf{r}}_0 + \dot{\mathbf{r}}_{ma} + \boldsymbol{\omega} \times \mathbf{r}_m. \tag{7.12-7}$$

Before using $(7\cdot12-7)$ for further discussion of the motion of a rigid body it will be of interest to apply it to the motion of a single particle. An obvious case of great significance in which the motion of a particle is referred to moving axes is that of a particle on the surface of the earth related to axes fixed in the earth and therefore rotating with the earth. To express the motion of such a particle with respect to a fixed system having the same origin, viz., the center of the earth, we must use eq. $(7\cdot12-7)$, noting that $\dot{\mathbf{r}}_0 \doteq 0$ in this case.

To write the equation of motion, we must differentiate (7·12-7)

with respect to the time. Thus the acceleration of the particle in the *fixed* system becomes

$$\ddot{\mathbf{r}}_f = \ddot{\mathbf{r}}_{ma} + \mathbf{\omega} \times \dot{\mathbf{r}}_m, \tag{7.12-8}$$

since $\dot{\omega}=0$, if we treat the angular velocity of the earth as approximately constant. We now use (7·12–4) in (7·12–8) and get

$$\ddot{\mathbf{r}}_f = \ddot{\mathbf{r}}_{ma} + \mathbf{\omega} \times \dot{\mathbf{r}}_{ma} + \mathbf{\omega} \times (\mathbf{\omega} \times \mathbf{r}_m). \tag{7.12-9}$$

But we must look into $\dot{\mathbf{r}}_{ma}$ more closely. From (7·12–3)

$$\dot{\mathbf{r}}_{ma} = \mathbf{i}_m \dot{x}_m + \mathbf{j} \dot{y}_m + \mathbf{k} \dot{z}_m. \tag{7.12-10}$$

Therefore

$$\ddot{\mathbf{r}}_{ma} = \mathbf{i}_m \ddot{\mathbf{x}}_m + \mathbf{j}_m \ddot{\mathbf{y}}_m + \mathbf{k}_m \ddot{\mathbf{z}}_m$$

$$+\dot{x}_m\frac{d\mathbf{i}_m}{dt}+\dot{y}_m\frac{d\mathbf{j}_m}{dt}+\dot{z}_m\frac{d\mathbf{k}_m}{dt}.$$
 (7·12–11)

But from (7.11–11) (which is perfectly general and not restricted to ${\bf M}$) this can be written

$$\ddot{\mathbf{r}}_{ma} = \mathbf{a}_m + \mathbf{\omega} \times \dot{\mathbf{r}}_{ma}, \qquad (7.12-12)$$

where \mathbf{a}_m is the apparent acceleration of the particle in the moving system, neglecting the rotation of the axes. Then (7.12-9) becomes

$$\ddot{\mathbf{r}}_f = \mathbf{a}_m + 2\mathbf{\omega} \times \dot{\mathbf{r}}_{ma} + \mathbf{\omega} \times (\mathbf{\omega} \times \mathbf{r}_m), \qquad (7.12-13)$$

and the corresponding resultant force on the particle of mass m in the fixed system becomes

$$\mathbf{F} = m\ddot{\mathbf{r}}_f = m\mathbf{a}_m + 2m\omega \times \dot{\mathbf{r}}_{ma} + m\omega \times (\omega \times \mathbf{r}_m). \quad (7.12-14)$$

We may rewrite this in terms of the apparent acceleration relative to the moving axes (i.e., those fixed in the earth) thus

$$\mathbf{a}_m = \mathbf{F}/m - 2\mathbf{\omega} \times \dot{\mathbf{r}}_{ma} - \mathbf{\omega} \times (\mathbf{\omega} \times \mathbf{r}_m).$$
 (7·12–15)

From its form it seems clear that the term $\omega \times (\omega \times r_m)$ corresponds to the well-known centripetal acceleration. On the other hand the term $2\omega \times \dot{r}_{ma}$ is relatively unfamiliar, though if we go back to eq. (3.6–9) we shall see that we have actually encountered it in connection with central field motion. There it is completely balanced by the acceleration due to angular acceleration. The acceleration $2\omega \times \dot{r}_{ma}$ is called the Coriolis acceleration. Its importance will be appreciated in a moment.

If the particle is subject only to the attraction of the earth we must write¹

$$\mathbf{F} = m\mathbf{g} + m\mathbf{\omega} \times (\mathbf{\omega} \times \mathbf{r}_m). \tag{7.12-16}$$

This combined with (7·12-15) yields

$$\mathbf{a}_m = \mathbf{g} - 2\mathbf{\omega} \times \dot{\mathbf{r}}_{m\alpha} \tag{7.12-17}$$

for the actual apparent acceleration relative to the earth. This is the quantity which is actually measured when we observe a falling body on the surface of the earth. If we write $(7\cdot12-17)$ in terms of the rectangular components we get

$$\ddot{x}_m = g_x + 2\omega \dot{y}_{ma},$$

$$\ddot{y}_m = g_y - 2\omega \dot{x}_{ma},$$

$$z_m = g_z,$$

$$(7.12-18)$$

recalling that $\omega_x = \omega_y = 0$, and $\omega_z = \omega$ in this special case. In Fig. 7.27 let us consider rectangular axes set up at the center of the earth O as indicated and let APB be one quarter of the trace of the earth's surface on the yz plane. Unit vectors \mathbf{i} , \mathbf{j} , \mathbf{k} are set up at P, and the horizon plane is indicated by WPN with N indicating the north, etc. PU is the zenith direction. The direction of the earth's axis is given by \mathbf{k} which makes angle λ with PN. Suppose a particle is dropped from rest at P in latitude λ . We then have $g_x = 0$, $g_y = -g \cos \lambda$, $g_z = -g \sin \lambda$. Let v be the observed speed of fall at any instant. Eqs. $(7\cdot12-18)$ then become

$$\ddot{x}_m = -2\omega v \cos \lambda,$$

$$\ddot{y}_m = -g \cos \lambda + 2\omega v \sin \lambda,$$

$$\ddot{z}_m = -g \sin \lambda.$$
(7·12–19)

The first equation shows that in addition to the acceleration due to gravity the falling particle also experiences an additional acceleration directed at right angles to the meridian plane in which it

¹ It should be emphasized that in $(7\cdot12-16)$ g is really defined in such a way as to agree with Newton's law of gravitation, i.e., $mg = -\frac{GmM}{R^2} r_{01}$, where M = mass of earth, R = radius of earth, and $r_{01} = \text{unit}$ vector. Moreover the particle is assumed to have no initial velocity.

starts to fall. This acceleration has the magnitude $2\omega v \cos \lambda$, is entirely a Coriolis effect, and is always directed toward the east as should be clear from the diagram.

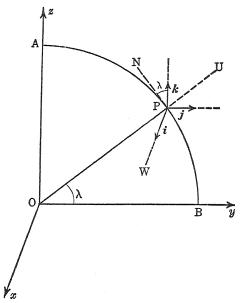


Fig. 7.27

Since v = gt, approximately, the magnitude of the easterly acceleration becomes

$$\left|\ddot{x}_{m}\right| = 2\omega gt \cos \lambda, \qquad (7.12-20)$$

and hence the total easterly deflection after time t from rest is approximately

$$x_m = \frac{\omega g t^3}{3} \cos \lambda = \frac{2}{3} \omega \sqrt{\frac{2h^3}{g}} \cos \lambda, \qquad (7.12-21)$$

if we use $h = \frac{1}{2}gt^2$ as the height from which the fall has taken place.

The Coriolis acceleration gives rise to cyclonic wind movements on the earth's surface.¹

¹ Cf. Page, "Introduction to Theoretical Physics" (D. Van Nostrand, New York, 2nd ed.), p. 107 ff.

7.13. Kinetic Energy of a Rigid Body. Let us revert to eq. (7.12-7) for the velocity of any point of a rigid body in the fixed set of axes, i.e., the axes fixed in space, and assume that one point of the body is fixed. The body then performs rotation about this point. It follows that both $\dot{r}_0 = 0$ and $\dot{r}_{ma} = 0$, the latter being true since the moving axes are fixed in the body. Hence (7.12-7) reduces to

$$\dot{\mathbf{r}}_f = \mathbf{\omega} \times \mathbf{r}_m \tag{7.13-1}$$

where \mathbf{r}_m is the position vector of the point P with respect to axes fixed in the body, and $\dot{\mathbf{r}}_f$ is the velocity of the corresponding point with respect to the axes fixed in space.

Let us now form the expression for the kinetic energy of the rigid body relative to the axes fixed in space. We have

$$K = \sum \frac{m_i}{2} \dot{\mathbf{r}}_{fi} \cdot \dot{\mathbf{r}}_{fi} = \sum \frac{m_i}{2} (\omega \times \mathbf{r}_{mi}) \cdot (\omega \times \mathbf{r}_{mi}), \quad (7.13-2)$$

wherein the sums are extended over all the particles of the body and the subscript i refers to any one particle. Now if we carry out the multiplication indicated in the right-hand term in $(7\cdot13-2)$ we can readily verify that

$$(\omega \times \mathbf{r}_{mi}) \cdot (\omega \times \mathbf{r}_{mi}) = \omega \cdot [\mathbf{r}_{mi} \times (\omega \times \mathbf{r}_{mi})]. \quad (7.13-3)$$

This is an important vector identity. In the present case it enables us to write (7.13-2) in the form

$$K = \frac{1}{2}\boldsymbol{\omega} \cdot \mathbf{M}, \qquad (7.13-4)$$

where, of course, we are introducing again the angular momentum **M** from $(7\cdot7-1)$. We should expect a connection between K in $(7\cdot13-4)$ and the work done by the resultant *torque* during the rotation of the rigid body, similar to that connecting the kinetic energy of translation of a particle and the work done by the resultant *force* during the translation [cf. eqs. $(1\cdot10-19, 20)$]. From $(7\cdot13-1)$ it follows that during time dt, the work done is

$$dW = \sum \mathbf{F}_i \cdot \dot{\mathbf{r}}_{fi} dt = \sum (\omega \times \mathbf{r}_{mi}) \cdot \mathbf{F}_i dt. \qquad (7.13-5)$$

But inspection shows that we can write $(\omega \times r_{mi}) \cdot F_i = \omega \cdot (r_{mi} \times F_i)$ and hence have in place of (7.13-5)

$$dW = \sum \mathbf{L}_i \cdot \mathbf{\omega} \, dt = \mathbf{L} \cdot \mathbf{\omega} \, dt, \tag{7.13-6}$$

where L is the resultant torque about the fixed point measured

with respect to axes fixed in the body. Now we revert to (7.7-5) and form the change in M, namely, dM, remembering that since the axes are fixed in the body the I_{xx} , etc., do not change. Thus we have

$$d\mathbf{M} = \mathbf{i}(I_{xx} d\omega_x - I_{xy} d\omega_y - I_{xz} d\omega_z)$$

$$+ \mathbf{j}(-I_{yx} d\omega_x + I_{yy} d\omega_y - I_{yz} d\omega_z)$$

$$+ \mathbf{k}(-I_{zx} d\omega_x - I_{zy} d\omega_y + I_{zz} d\omega_z). (7.13-7)$$

Let us form $\omega \cdot d\mathbf{M}$ with the result

$$\omega \cdot d\mathbf{M} = \omega_x (I_{xx} d\omega_x - I_{xy} d\omega_y - I_{xz} d\omega_z)$$

$$+ \omega_y (-I_{yx} d\omega_x + I_{yy} d\omega_y - I_{yz} d\omega_z)$$

$$+ \omega_z (-I_{zx} d\omega_x - I_{zy} d\omega_y + I_{zz} d\omega_z).$$

$$(7.13-8)$$

Next we write out the kinetic energy (7·13-4) as follows:

$$\frac{1}{2}\omega \cdot \mathbf{M} = \frac{1}{2}[\omega_x (I_{xx}\omega_x - I_{xy}\omega_y - I_{xz}\omega_z)
+ \omega_y (-I_{yx}\omega_x + I_{yy}\omega_y - I_{yz}\omega_z)
+ \omega_z (-I_{zx}\omega_x - I_{zy}\omega_y + I_{zz}\omega_z)].$$
(7.13-9)

Finally form the differential of $\frac{1}{2}\omega \cdot \mathbf{M}$. Thus

$$d(\frac{1}{2}\omega \cdot \mathbf{M}) = \omega_x (I_{xx} d\omega_x - I_{xy} d\omega_y - I_{xz} d\omega_z)$$

$$+ \omega_y (-I_{yx} d\omega_x + I_{yy} d\omega_y - I_{yz} d\omega_z)$$

$$+ \omega_z (-I_{zx} d\omega_x - I_{zy} d\omega_y + I_{zz} d\omega_z)$$

$$= \omega \cdot d\mathbf{M}, \qquad (7.13-10)$$

where we have, of course, utilized the fact that $I_{yx} = I_{xy}$, etc. The upshot is that

$$\mathbf{\omega} \cdot d\mathbf{M} = \mathbf{L} \cdot \mathbf{\omega} \, dt = dW = dK, \tag{7.13-11}$$

and therefore

$$W = K_1 - K_0 (7.13-12)$$

on integration between any two configurations of the body. This is the general work-kinetic energy relation for rotation of a rigid body with one point fixed.

The special case of zero resultant torque is of particular interest. From (7.7-2) there follows here

$$\mathbf{M} = \mathbf{Const.}, \qquad (7.13-13)$$

and hence from (7.13-10) we also have

$$2K = \omega \cdot \mathbf{M} = \text{Const.} \tag{7.13-14}$$

Consider now an angular velocity "space" in which every point corresponds to a set of values of ω_z , ω_y , ω_z . Since from (7·13–9) twice the kinetic energy becomes

$$\mathbf{\omega} \cdot \mathbf{M} = I_{xx}\omega_x^2 + I_{yy}\omega_y^2 + I_{zz}\omega_z^2 - 2I_{xy}\omega_x\omega_y - 2I_{yz}\omega_y\omega_z - 2I_{zz}\omega_z\omega_x, \quad (7.13-15)$$

it follows that (7·13–14) is the equation of a quadric surface in this space. By proper choice of the axes fixed in the body the products of inertia may be made to vanish¹ and the equation of the quadric surface becomes

$$I_{xx}\omega_x^2 + I_{yy}\omega_y^2 + I_{zz}\omega_z^2 = 2K.$$
 (7·13–16)

This is the equation of an ellipsoid in the angular velocity space with its center at the origin. The magnitude of the vector from the origin of coördinates in the ω space to any point on the ellipsoid (7·13–16) represents a possible value of the angular speed consistent with the given kinetic energy. The ellipsoid has been called the *momental* ellipsoid or ellipsoid of Poinsot. It is clear that in the direction corresponding to the minor axis of the ellipsoid (the principal axis corresponding to maximum moment of inertia) the angular speed is least, whereas in the direction of the major axis the angular speed is greatest.

The change in ω in passing from one point of the momental ellipsoid to another is given by

$$d\mathbf{\omega} \cdot \mathbf{M} = 0. \tag{7.13-17}$$

Now $d\omega$ must lie in the plane tangent to the ellipsoid at the point in question. Hence the tangent plane is perpendicular to M. But M is a constant vector since we are still considering the resultant torque to be zero. Therefore in this case the tangent

¹ Cf., for example, Smith and Gale, "New Analytic Geometry" (Ginn & Co., Boston, 1912), p. 317 ff.

plane remains fixed in position relative to a set of fixed axes. If we set P = W/2V

 $\mathbf{p} = \mathbf{M}/2K, \qquad (7.13-18)$

the vector \mathbf{p} is normal to the tangent plane, and eq. (7.13-14) becomes

 $\boldsymbol{\omega} \cdot \mathbf{p} = 1. \tag{7.13-19}$

But $\omega \cdot \mathbf{p}$ is p times the projection of ω along the normal to the tangent plane. It follows that the tangent plane remains at a constant distance from the origin as the motion goes on. The motion of the rigid body then can be represented by a rolling (without slipping) of the momental ellipsoid on the fixed tangent plane. During the rolling the vector ω traces out a cone relative to the axes fixed in the body with vertex at the fixed point. This cone intersects the momental ellipsoid in a curve called by Poinsot the polhode. It intersects the invariable tangent plane in a curve called the herpolhode. For illustrations of these curves, A. G. Webster's "Dynamics" may be consulted.

7.14. Euler's Equations of Motion. We have not yet exhausted the utility of axes fixed in the rigid body. We shall continue to restrict our attention to the case in which one point is fixed and take this as the origin of rectangular coördinates fixed in the body. However, we shall now insist that the axes are so chosen that the products of inertia I_{xx} , etc., all vanish, leaving only the moments of inertia I_{xx} , etc., in the expression for M in (7.7-5). Axes for which this is true are termed the principal axes and there exist mathematical methods for finding them in each case. If the body is symmetrical about the fixed point, the three mutually perpendicular axes of symmetry through this point will be principal axes. We have then

$$\mathbf{M} = \mathbf{i}\omega_x I_{xx} + \mathbf{j}\omega_y I_{yy} + \mathbf{k}\omega_z I_{zz}. \tag{7.14-1}$$

Now from (7·11-11) we have

$$\dot{\mathbf{M}} = \mathbf{i} I_{xx} \dot{\omega}_x + \mathbf{j} I_{yy} \dot{\omega}_y + \mathbf{k} I_{zz} \dot{\omega}_z
+ \omega \times \mathbf{M}
= \mathbf{i} [I_{xx} \dot{\omega}_x + (I_{zz} - I_{yy}) \omega_y \omega_z]
+ \mathbf{j} [I_{yy} \dot{\omega}_y + (I_{xx} - I_{zz}) \omega_z \omega_x]
+ \mathbf{k} [I_{zz} \dot{\omega}_z + (I_{yy} - I_{xz}) \omega_x \omega_y].$$
(7.14-2)

The fundamental equation of rotational motion (7.7-2) then takes a particularly simple form and in terms of rectangular components becomes the three equations

$$\begin{split} I_{xx}\dot{\omega}_x + (I_{zz} - I_{yy})\omega_y\omega_z &= L_x, \\ I_{yy}\dot{\omega}_y + (I_{xx} - I_{zz})\omega_z\omega_x &= L_y, \\ I_{zz}\dot{\omega}_z + (I_{yy} - I_{xx})\omega_x\omega_y &= L_z. \end{split} \tag{7.14-3}$$

These are known as Euler's equations of motion of the rigid body with one point fixed.

Let us specialize to the case in which the resultant torque vanishes, i.e., $L_x = L_y = L_z = 0$. Then multiply through the resulting equations by ω_x , ω_y , ω_z , respectively, and add. The algebra gives

$$\frac{d}{dt}\left[I_{xx}\omega_{x^{2}}+I_{yy}\omega_{y^{2}}+I_{zz}\omega_{z^{2}}\right]=0,$$

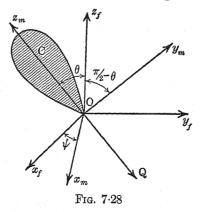
or

$$I_{xx}\omega_{x}^{2} + I_{yy}\omega_{y}^{2} + I_{zz}\omega_{z}^{2} = const.$$
 (7.14-4)

From (7·13-15) the expression on the left is simply twice the kinetic energy of the body (recalling that the products of inertia are now zero) and we have again the fact that the kinetic energy remains constant under the action of zero resultant torque. It is appropriate to point out at this place that the constancy of the kinetic energy does not necessarily imply that there is no angular

acceleration, since in $(7\cdot14-4)$ ω_z , ω_y , and ω_z may change with the time without invalidating the equation.

7.15. The Motion of a Top. This is probably the most interesting elementary application of Euler's equations. The top is assumed to be a solid of revolution with one end of the axis of symmetry (i.e., the peg) fixed in space. This axis is also taken as the spin axis. The



fixed point (cf. Fig. 7.28) is taken as the origin O of a set of axes fixed in space x_f , y_f , z_f . With O as origin we also set up another

set of axes x_m , y_m , z_m in which z_m is fixed to the body along its axis of symmetry. Ordinarily we should also expect to fix the axes x_m and y_m to the body, but it now proves more convenient to let x_m and y_m rotate with angular velocity ω which is different from that of spin, which we denote by s. This will not affect the application of our previous reasoning since the moments of inertia about all axes perpendicular to the axis of spin are the same (i.e., $I_{xx} = I_{yy}$) and do not change with the time as the axes x_m and y_m rotate. We choose as the x_m axis the line of intersection of the plane through O perpendicular to the z_m axis and the $x_f y_f$ plane. Let the angle between x_m and x_f be ψ and that between z_m and z_f be θ . These are usually known as the *Eulerian* angles. There is a third, namely, ϕ , which a line OQ in the $x_m y_m$ plane rotating with the top makes with the x_m axis. The components ω_x , ω_y , ω_z about the x_m , y_m , and z_m axes can be expressed in terms of $\dot{\theta}$ and $\dot{\psi}$ as follows:

$$\omega_x = \hat{\theta},
\omega_y = \dot{\psi} \sin \theta,
\omega_z = \dot{\psi} \cos \theta.$$
(7.15-1)

Note that the x_m axis always lies in the $x_f y_f$ plane. This accounts for the expression for ω_x . Note also that z_f , z_m , and y_m are all in the same plane. As the figure indicates, ψ is the angular velocity of the rotating axes about the z_f axis. It is known as the *precessional* velocity of the top.

The resultant torque on the top is due to gravity. If we denote the mass as m and locate it as usual at the center of mass C (with OC = l) the component torque about the x_m axis is $mgl \sin \theta$. The components about the y_m and z_m axes vanish. Before we can write Euler's equations we must rewrite M to take account of the spin velocity s. Thus we now write (recalling that $I_{xx} = I_{yy}$)

$$\mathbf{M} = \mathbf{i} I_{xx} \omega_x + \mathbf{j} I_{xx} \omega_y + \mathbf{k} I_{zz} (\omega_z + s), \qquad (7.15-2)$$

whence

$$\dot{\mathbf{M}} = \mathbf{i} I_{xx} \dot{\omega}_x + \mathbf{j} I_{xx} \dot{\omega}_y + \mathbf{k} I_{zz} (\dot{\omega}_z + s) + \mathbf{i} [\omega_y I_{zz} (\omega_z + s) - \omega_z I_{xx} \omega_y] + \mathbf{j} [\omega_z I_{xx} \omega_x - \omega_x I_{zz} (\omega_z + s)] + \mathbf{k} [\omega_x I_{xx} \omega_y - \omega_y I_{xx} \omega_x].$$
 (7.15-3)

Hence Euler's equations (7.14-3) now take the form

$$I_{zz}\dot{\omega}_{x} + (I_{zz} - I_{xx})\omega_{y}\omega_{z} + I_{zz}\omega_{y}s = mgl\sin\theta,$$

$$I_{xx}\dot{\omega}_{y} + (I_{xx} - I_{zz})\omega_{x}\omega_{z} - I_{zz}\omega_{x}s = 0,$$

$$I_{zz}(\dot{\omega}_{z} + \dot{s}) = 0.$$
(7.15-4)

These can be somewhat simplified by writing $s + \omega_z = S$, so that

$$\begin{split} I_{xx}\dot{\omega}_{x} - I_{xx}\omega_{y}\omega_{z} + SI_{zz}\omega_{y} &= mgl\sin\theta, \\ I_{xx}\dot{\omega}_{y} + I_{xx}\omega_{x}\omega_{z} - SI_{zz}\omega_{x} &= 0, \\ I_{zz}\dot{S} &= 0. \end{split} \tag{7.15-5}$$

The last equation says that the resultant angular speed about the z_m axis is constant. If we multiply the first equation by ω_x , the second by ω_y , and the third by S and add, we obtain, after integrating,

$$\frac{1}{2}I_{xx}(\omega_x^2 + \omega_y^2) + \frac{1}{2}I_{zz}S^2 + mgl\cos\theta = E, \quad (7.15-6)$$

where E is the total energy of the motion, $mgl\cos\theta$ is the potential energy with respect to the horizontal plane, and the rest of the left-hand side is the resultant kinetic energy. Eq. (7·15–6) is then the energy equation of the top.

A special case of the first equation in (7.15-5) is not without interest. Suppose that the spin speed s is so large that the term $SI_{zz}\omega_y$ dominates the right-hand side. Then $I_{xz}\dot{\omega}_x - I_{xx}\omega_y\omega_z$ may be neglected and the equation becomes

$$SI_{zz}\dot{\psi} = mgl, \qquad (7.15-7)$$

which says that the precessional velocity is inversely proportional to the spin velocity and the moment of inertia about the spin axis. This is the familiar result of the elementary theory of the simple gyroscope.

To treat the problem more generally we note that the torque about the z_f axis is zero and hence the angular momentum about this axis is constant. But the latter is the sum of the components of **M** along z_f . We therefore have

$$I_{zz}S\cos\theta + I_{xx}\omega_y\sin\theta = A = \text{constant.}$$
 (7.15–8)

Written in terms of the precessional velocity $\dot{\psi}$, this yields

$$\dot{\psi} = \frac{A - I_{zz}S\cos\theta}{I_{zz}\sin^2\theta}.$$
 (7.15–9)

Moreover the energy equation (7.15-6) can be written

$$\dot{\theta}^2 + \psi^2 \sin^2 \theta = \frac{2E - 2mgl \cos \theta - I_{zz}S^2}{I_{xx}}$$
 (7·15–10)

With the introduction of the new quantities

$$\alpha = A/I_{xx}, \quad \beta = I_{zx}S/I_{xx},$$

$$a = \frac{2E - I_{zz}S^2}{I_{xx}}, \quad b = 2mgl/I_{xx},$$

$$(7.15-11)$$

eqs. (7.15-9) and (7.15-10) take on the simpler form

$$\dot{\psi} = \frac{\alpha - \beta \cos \theta}{\sin^2 \theta},$$

$$\dot{\theta}^2 + \dot{\psi}^2 \sin^2 \theta = \alpha - b \cos \theta.$$
(7.15–12)

The elimination of ψ between these equations yields

$$\dot{\theta} = -\frac{(\alpha - \beta \cos \theta)^2}{\sin^2 \theta} + a - b \cos \theta. \qquad (7.15-13)$$

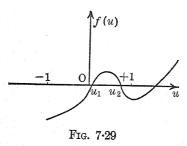
Letting $u = \cos \theta$, we get

$$\dot{u}^2 = (a - bu)(1 - u^2) - (\alpha - \beta u)^2 = f(u). \quad (7.15-14)$$

Evidently the variation of θ with time depends on the quadrature

$$t = \int \frac{du}{\sqrt{f(u)}} + \text{constant}, \qquad (7.15-15)$$

which in turn can be expressed only in terms of elliptic functions. However, we can learn something about the top's behavior with-



out carrying out the integration in $(7\cdot15-15)$. Clearly f(u) must be positive to insure the reality of u, and, further, u must lie between 0 and +1. The function f(u) is a cubic with three roots of which two must lie between 0 and +1. The situation is depicted in Fig. 7·29. We shall call the roots of f(u) lying in the

allowed interval u_1 and u_2 . It follows that the inclination of the top's axis to the vertical is restricted to lie within the limiting

angles $\theta_1 = \arccos u_1$ and $\theta_2 = \arccos u_2$. More elaborate consideration of the problem shows that θ varies periodically between these limits. This motion is known as *nutation*. In general the spinning of a top or gyroscope is accompanied by both precession and nutation. Under certain conditions the nutation vanishes: this demands, of course, that $\theta_1 = \theta_2$ always. Hence $u_1 = u_2$ must be a double root of f(u), and for vanishing nutation we must be able to write this function in the form

$$f(u) = B(u - u_3)(u - u_1)^2,$$
 (7·15-16)

where B is constant. It is clear that $u = u_1$ must satisfy both f(u) = 0 and df/du = 0. This will suffice to fix u_1 , and the precessional velocity ψ , given by the first equation in $(7\cdot15-12)$, becomes in terms of u_1

$$\dot{\psi} = \frac{\beta}{2u_1} [1 \pm \sqrt{1 - 2bu_1/\beta^2}]. \tag{7.15-17}$$

For each u_1 there are then two possible precessional velocities.

If the spin velocity is so great that $\beta^2 \gg 2bu_1$, the two values of ψ are

$$\dot{\psi}_1 = mgl/I_{zz}S, \qquad (7.15-18)$$

$$\dot{\psi}_2 = \frac{I_{zz}S}{u_1 I_{xx}} \left[1 - \frac{mglu_1 I_{xx}}{I_{zz}^2 S^2} \right]. \tag{7.15-19}$$

Evidently $\psi_1 \ll \psi_2$. The slower precession is usually the one observed.

Suppose the top is started with axis vertical and S = s, i.e., with spin velocity only. This leads at once to $\alpha = \beta$ and $\alpha = b$. The roots of f(u) then become

$$u_1 = +1$$
, $u_2 = +1$, $u_3 = \alpha^2/a - 1$.

As long as $S^2 > 4mglI_{xx}/I_{zz}^2$, $u_3 > 1$, and the motion of the top is confined to simple rotation about the vertical axis with no precession or nutation, i.e., the top sleeps. As a result of friction at the point of support, energy may be lost until $S^2 < 4mglI_{xx}/I_{zz}^2$ when u_3 becomes less than unity, and precession sets in with nutation between the angles $\theta = 0$ and $\theta = \theta_2$, where

$$\cos \theta_2 = I_{zz}^2 S^2 / 2mg II_{xx} - 1. \tag{7.15-20}$$

This means the top begins to wobble. With the continued decrease of S, θ_2 increases until the top falls down.

PROBLEMS

1. A rigid body possesses velocity of translation v and rotational velocity ω . What is the velocity of a point P of the rigid body with position vector r with respect to a fixed origin? If v and ω are constant in time, what is the path of the point P and what is its equation?

2. A flat cylindrical disc of brass 10 cm in diameter and 2 cm in thickness has coiled about its edge a light flexible and inextensible string. If the free end of the string is attached to a fixed point and the disc is allowed to fall, what

will be its total kinetic energy at the end of 2 seconds?



3. A steel ring in the form of a torus with inside diameter 50 cm and outside diameter 52 cm is allowed to perform small oscillations in a vertical plane about a rod on which the ring is hung. (See figure.) Calculate the frequency of the vibrations. How much torque is needed to displace the ring through an angle of 5° from its equilibrium position?

4. An elliptical disc of brass of major and minor axes 20 cm and 10 cm respectively and 2 cm thick may rotate about an axis through one focus and perpendicular to the disc. If the disc is displaced so that its major axis is horizontal and let go, find the angular velocity of the disc at the instant when the major axis is vertical.

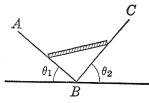
5. Consider two particles of mass m joined by a weightless rod of length l. The system is assumed to rotate about an axis through the center of the rod and perpendicular to it. The quantum theory requires that only those rotational motions are possible for which the angular integral of the angular momentum over the period from 0 to 2π (viz. $\int_0^{2\pi} M d\theta$) is equal to some multiple of Planck's constant h. Find the expression for the allowed values of the energy of rotation and compare with the similar expressions for the quantized energy states of a hydrogen atom (Sec. 4.4).

- 6. A non-homogeneous circular rod has line density varying with distance x along it of the form $\rho = \rho_0 + kx$. The length of the rod is l. The rod falls from rest in a vertical position to a horizontal floor. If the bottom in contact with the floor does not slip, with what kinetic energy will it strike? What will be the velocity of the top end?
- 7. Calculate the translational acceleration of a sphere rolling down an inclined plane of angle θ . First solve the problem by the torque equation and then by energy considerations.
- 8. The ring of Problem 3 rolls down a 30° incline. How much velocity does it gain in 1 second?

- 9. Find the period of the small oscillations of a cube of side a about one edge and show that the length of the equivalent simple pendulum is $\frac{2\sqrt{2}}{3}a$. Also find the center of oscillation.
- 10. Show that the moment of inertia of a thin rectangular sheet with sides a and b about an axis through the center and parallel to the side b is $ma^2/12$, where m is the mass of the sheet.
- 11. Prove that the moment of inertia of a homogeneous spherical shell with internal and external radii r_1 and r_2 respectively about any diameter is $\frac{2}{5} m \frac{r_2{}^5 r_1{}^5}{r_2{}^3 r_1{}^3}.$
- 12. Prove that the moment of inertia of a solid homogeneous sphere about an axis tangent to the sphere is $\frac{7}{5} \cdot ma^2$, where m is the mass and a the radius of the sphere.
- 13. Prove that the moment of inertia of a homogeneous plane triangular plate about any axis in its plane is equal to the moment of inertia about the same axis of three masses each placed at a midpoint of a side of the triangle and each equal to one third the mass of the plate.
- 14. Consider a non-homogeneous ellipsoid with semi-axes a, b, c respectively and with density varying directly as the distance from the center along the longest axis (the layers of equal density being concentric elliptical sheets perpendicular to the longest axis). Show that the moment of inertia of the ellipsoid with respect to its longest axis is $\frac{2}{9} m(b^2 + c^2)$.
- 15. Prove that the product of inertia of a rigid body with respect to any two mutually perpendicular rectangular coördinate axes is equal to the product of inertia with respect to two parallel axes through the center of mass plus the product of inertia of the whole mass of the body located at the center of mass with respect to the original axes.
- 16. A uniform wire is bent into the form of a triangle. Find the position of the center of mass.
- 17. Out of a uniform circular disc of radius a, one quadrant is cut. Find the center of mass of the remainder.
- 18. A uniform flexible cord is suspended between two points in the same horizontal line. Calculate the position of the center of mass of the cord.
- 19. In a hemisphere of radius a the density varies inversely as the distance from the center. Determine the position of the center of mass.
 - 20. Find the center of mass of one octant of a homogeneous sphere.
- 21. Determine the coördinates of the center of mass of one octant of the ellipsoid $x^2/a^2 + y^2/b^2 + z^2/c^2 = 1$, in which the density is a linear function of the distance along the x axis.

- 22. Prove that the moment of a couple has the same value about any axis perpendicular to the plane of the couple.
- 23. Show that a couple may be replaced by any other couple with the same moment in the same plane.

24. A uniform rod of mass m and length l rests on two inclined planes AB and BC (see figure), with their line of intersection lying in a horizontal plane.



The rod lies in a vertical plane perpendicular to the line of intersection. The angles of the planes are θ_1 and θ_2 respectively. Find the position of equilibrium of the rod if its contact with both planes is smooth. Calculate the normal thrusts. What would be the effect on the foregoing result if one plane only is smooth and the other rough?

- 25. A uniform ladder rests with one end against a rough horizontal plane and the other end against an equally rough vertical plane. What is the smallest coefficient of friction that will allow the ladder to rest in all positions?
- 26. A uniform rod of length l and mass m rests against the horizontal rim of a hemispherical bowl of radius a, its lower end touching the inside of the bowl. Find the position it will assume for equilibrium.
- 27. Given a circular table with three legs vertically below the rim and forming an equilateral triangle. Find the smallest weight which when placed on the table is able to upset it.
- 28. Write the equation of motion of a general rigid body which rotates about the fixed x axis and show that in general in addition to the torque component about the x axis (due to gravity, for example) there must be torque components about the y and z axes as well. What is the nature of the forces producing the latter torques?
- 29. Calculate the magnitude of the easterly deflection from the plumb line experienced by a particle dropped from rest at a height of 500 feet in latitude 45° N.
- 30. A railroad train of mass 100 tons moves due east in latitude 45° N with constant velocity of 60 mi/hr. Find the magnitude of the force at right angles to the direction of motion. How is it directed?
- 31. A particle is thrown vertically upward in a vacuum with velocity \mathbf{v} . Prove that on its return it hits the ground a distance $\frac{4\omega v^3}{3g^2} \cdot \cos \lambda$ west of the place from which it was thrown. (The latitude is λ and ω is the angular velocity of the earth.)
- 32. A simple pendulum has its point of suspension directly over the center of a horizontal turntable rotating with angular velocity ω . If the bob is pulled aside and let go, what sort of curve will it trace out on the turntable?

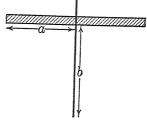
33. A bullet is fired horizontally with a velocity of 1000 meters/sec in latitude 45°N. What is the deflection from the plane of projection in direction and magnitude after 1 second of flight?

34. Find the equation of the momental ellipsoid of a homogeneous elliptical plate in the xy plane with semi-major and semi-minor axes a and b respectively.

35. Prove that the momental ellipsoid of any regular polyhedron with its center at the origin of rectangular coördinates is a sphere. Specialize to the case of a cube and find the radius of the corresponding sphere.

36. An airplane motor and propeller has a moment of inertia about the axis of spin of 25 kilogram meters². If it is moving instantaneously in a curved path with a radius of curvature of 150 meters with a velocity of 300 km/hr, what is the torque (direction and magnitude) tending to make the plane move vertically? Where does the counter torque

37. A simple top may be formed by piercing the center of a uniform disc of radius a and mass m. We may assume that the mass of the pin is negligible compared with the mass of the disc. Calculate the moment of inertia about the mutually perpendicular axes passing through O, the point of contact of the pin with the floor, the pin itself being



taken as the z axis. Compute the minimum spin velocity at which the top will "sleep." At what spin velocity will the top begin to roll on the ground?

38. The equation of motion for an electron of mass m and charge e moving in an electric field \mathbf{F} and magnetic field \mathbf{H} is in vector form

$$ma = e\mathbf{F} + \frac{e}{c}\mathbf{v} \times \mathbf{H}$$

where ${\bf v}$ is the velocity of the electron and c is numerically equal to the velocity of light. The first term on the right is the force due to the electric field, whereas the second is that due to the magnetic field. (See, for example, Page and Adams, *Principles of Electricity*, D. Van Nostrand Co., N. Y., 1931, Chap. VIII.) Find the form assumed by this equation when the motion of the electron is referred to a system of axes rotating about the direction of the

magnetic field with constant angular velocity $\omega = -\frac{e}{2mc}H$. In particular show that if terms proportional to H^2 are neglected the equation of motion in the moving system reduces to the form

 $ma_m = eF$.

This is known as Larmor's theorem, and the angular velocity of the rotating axes is called the Larmor precession. It is of vital importance in the study of the effect of a magnetic field on an atom.

39. Using the appropriate values for e, m, and c, calculate the value of the Larmor precession in the preceding problem for a magnetic field of 10,000 gauss. Consider an electron moving in the circular orbit of smallest quantum number (viz., $n_1 + n_2 = 1$) in the Bohr model of the hydrogen atom (see Sec. 3.9 and Fig. 3.9). Calculate the numerical magnitudes of the Coriolis force, viz., $2m\omega \times v_a$ and the centripetal force, $m\omega \times (\omega \times r_m)$, acting on the electron in the axes rotating with the above precessional velocity ω . Compare the latter with the actual centripetal force on the electron in its circular orbit about the nucleus, unaffected by the field.

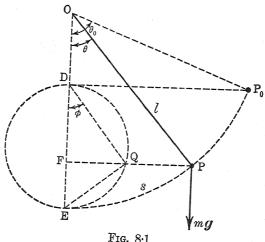




CHAPTER VIII

CONSTRAINED MOTION

8.1. Simple Types of Constraints. The Simple Pendulum. There are many important cases of motion of a particle in which the latter is compelled by the geometry of its environment to move on some specified curve or surface. We have already encountered in Chapter II a simple illustration in motion on an inclined plane. Another example is furnished by the simple pendulum, where the motion of the bob must take place along the



arc of a circle due to the invariable length of the string. We have indeed taken up already the approximate case of small motions in connection with the physical pendulum. This will be a good place to investigate the more general case where the swing may be of arbitrary amplitude. Consider Fig. 8-1 where the bob P, which for our purposes here is to be considered a single particle, is suspended by a string of length l from the fixed point O. The bob is in equilibrium under gravity when OP is vertical, i.e., coincides with OE. We wish to determine the motion resulting

when the bob is pulled aside through any arbitrary arc (less than $\pi/2$ for our convenience) and then let go.

The force on the bob in the direction of the motion is of magnitude $mg \sin \theta$. If we denote the displacement along the arc by s, the equation of motion is

$$m\ddot{s} = -mg\sin\theta. \tag{8.1-1}$$

Now $s = l\theta$ and hence (8.1-1) becomes

$$\ddot{\theta} = -\frac{g}{l}\sin\theta. \tag{8.1-2}$$

A first integration may be carried out by multiplying both sides by $\dot{\theta} dt$, whence

$$\frac{1}{2}\dot{\theta}^2 = \frac{g}{l}\cos\theta + C_1. \tag{8.1-3}$$

The constant C_1 may be evaluated by the initial condition that $\dot{\theta} = \text{zero for a value of } \theta = \theta_0$, the amplitude angle. Then there results

$$\frac{1}{2}\dot{\theta}^2 = \frac{g}{l}\left(\cos\theta - \cos\theta_0\right). \tag{8.1-4}$$

Now if the initial angle θ_0 is small (say < 10°) we may write

$$\cos\theta = 1 - \frac{\theta^2}{2},\tag{8.1-5}$$

and on substitution and rearrangement

$$\frac{d\theta}{\sqrt{\theta_0^2 - \theta^2}} = \sqrt{\frac{g}{l}} dt. \tag{8.1-6}$$

Integration yields

$$\arcsin\frac{\theta}{\theta_0} = \sqrt{\frac{g}{l}}t + C_2,$$

or

$$\theta = \theta_0 \sin\left(\sqrt{\frac{g}{l}}t + C_2\right). \tag{8.1-7}$$

In this case the motion is simple harmonic with frequency

$$\nu = \frac{1}{2\pi} \sqrt{\frac{g}{l}}.$$
 (8·1-8)

Look now at the motion in the case of arbitrary amplitude; we have the general integral

$$\frac{d\theta}{\sqrt{\cos\theta - \cos\theta_0}} = \sqrt{\frac{2g}{l}} dt. \tag{8.1-9}$$

From elementary trigonometry

$$\cos\theta = 1 - 2\sin^2\left(\frac{\theta}{2}\right).$$

On substitution into (8·1-9) we have therefore

$$\frac{d\left(\frac{\theta}{2}\right)}{\sqrt{\sin^2\left(\frac{\theta_0}{2}\right) - \sin^2\left(\frac{\theta}{2}\right)}} = \sqrt{\frac{g}{l}} dt.$$
 (8·1-10)

We shall now introduce a new angle ϕ defined by the relation

$$\sin\left(\frac{\theta}{2}\right) = \sin\left(\frac{\theta_0}{2}\right) \cdot \sin\phi = k \sin\phi,$$

whence

$$\frac{\theta}{2} = \arcsin (k \sin \phi),$$

so that

$$d\left(\frac{\theta}{2}\right) = \frac{k\cos\phi\,d\phi}{\sqrt{1 - k^2\sin^2\phi}}.$$

We can see the physical meaning of the angle ϕ by letting the initial position of P be P_0 (Fig. 8·1), so that $\angle EOP_0 = \theta_0$. Then draw P_0D normal to DE and construct the circle with diameter ED. Let the perpendicular PF from P to ED cut this circle at Q. Then the angle $EDQ = \phi$. For

$$\overline{EF} = l(1 - \cos \theta) = 2l \sin^2 \left(\frac{\theta}{2}\right),$$

and assuming that \angle EDQ really is ϕ , we have

$$\frac{\overline{EQ}}{\overline{ED}} = \sin \phi = \frac{\overline{EF}}{\overline{EQ}},$$

so that

$$\overline{EF} = \overline{ED} \sin^2 \phi.$$

which from Fig. 8.1 becomes

$$\overline{EF} = 2l \sin^2\left(\frac{\theta_0}{2}\right) \cdot \sin^2\phi.$$

Hence $\sin\left(\frac{\theta}{2}\right) = \sin\left(\frac{\theta_0}{2}\right) \cdot \sin \phi$, as above. Substitution into (8·1–10) now yields

$$\frac{d\phi}{\sqrt{1-k^2\sin^2\phi}} = \sqrt{\frac{g}{l}}\,dt,\tag{8.1-11}$$

and on integration we have

$$t = \sqrt{\frac{l}{g}} \int_0^{\phi} \frac{d\phi}{\sqrt{1 - k^2 \sin^2 \phi}},$$
 (8·1-12)

where t is the time taken by the particle in moving between the positions $\theta = \theta$ and $\theta = 0$. The integral appearing here is the well known elliptic integral $F(k, \phi)$ connected with the elliptic functions. Its values are tabulated in Peirce's Table of Integrals. Let us note a few special cases. For example, take $\theta_0 = 10^{\circ}$ and calculate the quarter period, i.e., let $\phi = \pi/2$. Then k = .0872 and we have, denoting the period by P,

$$P = 6.2952 \sqrt{\frac{l}{g}}.$$

To the same number of places, $2\pi = 6.2832$. This indicates that the simple formula $(8\cdot 1-8)$ is accurate for amplitude angles less than 10° with an error of less than .2%. For $\theta_0 = 20^\circ$ we have k = .1736 and $P = 6.3312 \sqrt{l/g}$. The error for a single period is still less than 1%. For many practical purposes the integrand in $(8\cdot 1-12)$ may be expanded in a series and the integration carried out term by term. Thus expanding $(1 - k^2 \sin^2 \phi)^{-\frac{1}{2}}$,

$$t = \sqrt{\frac{l}{g}} \int_0^{\phi} (1 + \frac{1}{2}k^2 \sin^2 \phi + \cdots) d\phi$$

$$= \sqrt{\frac{l}{g}} \left\{ \phi + \frac{k^2}{4} (\phi - \frac{1}{2} \sin 2\phi) + \cdots \right\}. \quad (8.1-13)$$

¹ See E. B. Wilson, "Advanced Calculus" (Ginn & Co., Boston, 1911), p. 503.

Hence for the period P we have

$$P = 2\pi \sqrt{\frac{l}{g}} \left[1 + \frac{k^2}{4} + \cdots \right] = 2\pi \sqrt{\frac{l}{g}} \left(1 + \frac{\theta_0^2}{16} + \cdots \right), \quad (8.1-14)$$

if we recall that $k = \sin \theta_0/2 = \theta_0/2$, if θ_0 is small. The formula $(8\cdot 1-14)$ is frequently very useful for small initial amplitudes.

8.2. Motion of a Particle on a Smooth Surface of Arbitrary Form. Let us suppose that the equation of the surface on which a particle is constrained to move is given by

$$\phi(x, y, z) = 0. (8.2-1)$$

The particle (of mass m) is acted on by the external force \mathbf{F} with components F_z , F_y , F_z , while the surface exerts on it a reaction force \mathbf{R} , with components R_z , R_y , R_z . Now if the surface is smooth the force \mathbf{R} acts along the *normal* and hence

$$R_x = \lambda R, \quad R_y = \mu R, \quad R_z = \nu R, \tag{8.2-2}$$

where λ , μ , ν are the direction cosines of the normal drawn in the direction of R, and will be functions of x, y, and z in general, being constant only for the case of a plane. The component equations of motion of the particle are then

$$m\ddot{x} = F_x + \lambda R,$$

 $m\ddot{y} = F_y + \mu R,$
 $m\ddot{z} = F_z + \nu R.$ (8.2-3)

Let us multiply these equations through by \dot{x} , \dot{y} , \dot{z} respectively and add. We then get

$$\begin{split} \frac{m}{2} \, \frac{d(\dot{x}^2)}{dt} + \frac{m}{2} \, \frac{d(\dot{y}^2)}{dt} + \frac{m}{2} \, \frac{d(\dot{z}^2)}{dt} &= \frac{1}{2} \, m \, \frac{d(v^2)}{dt} \\ &= \dot{x} F_x + \dot{y} F_y + \dot{z} F_z \\ &+ R(\dot{x}\lambda + \dot{y}\mu + \dot{z}\nu). \end{split} \tag{8.2-4}$$

Now \dot{x} , \dot{y} , \dot{z} are proportional at any point to the direction cosines of the tangent to the surface at that point since the particle is

¹ For further discussion of the pendulum, reference may be made to P. G. Tait and W. J. Steele, "Dynamics of a Particle" (Macmillan, 7th ed., London, 1900), Chap. VI.

constrained to remain on the surface. Hence $\dot{x}\lambda + y\mu + \dot{z}\nu$ is proportional to the cosine of the angle between the tangent and the normal to the surface and is thus zero. Eq. (8·2–4) may then be written

$$\frac{1}{2}md(v^2) = F_x dx + F_y dy + F_z dz.$$
 (8.2–5)

Integration yields

$$\frac{1}{2}mv^2 = \int (F_x dx + F_y dy + F_z dz) + C. \qquad (8.2-6)$$

If the external forces involved are conservative (Sec. 4·1), we may write the integral as -V(x, y, z) and have

$$\frac{1}{2}mv^2 + V(x, y, z) = C, (8.2-7)$$

where C is a constant. This is the energy equation and we note the interesting fact that as might have been anticipated the constraints do not enter it. The actual path followed by the particle can obviously not be determined by the second integration of $(8\cdot2-7)$ alone for this furnishes but one of the necessary equations. As long as the time is involved the number of the latter is three (i.e., we must express x, y, z in terms of t). We get the remaining two by eliminating R among the eqs. $(8\cdot2-3)$. These yield

$$\frac{m\ddot{x} - F_x}{\lambda} = \frac{m\ddot{y} - F_y}{\mu} = \frac{m\ddot{z} - F_z}{\nu}.$$
 (8·2–8)

On integrating these two equations we get two relations involving x, y, z, and t. Elimination of the time yields a relation among x, y, and z, which represents a second surface. The intersection of this with the surface $\phi(x, y, z) = 0$ is the path of the particle, while the equation (8·2–7) can then be used to give the "time equation," that is, the equation which expresses the rate at which the motion takes place along the path. This is, in formal outline, the general method for solving a problem in constrained motion on a smooth surface.

As an illustration of the preceding general method let us consider some aspects of the motion of a material particle under the influence of gravity, but constrained to move in a smooth hemispherical bowl. For the sake of simplicity we shall choose the z axis vertical and directed downward, and take the origin at the

center of the sphere. At the point P(x, y, z) the direction cosines of the *inward* drawn normal (R is directed in) are (Fig. 8.2)

$$\lambda = -\frac{x}{a}, \quad \mu = -\frac{y}{a}, \quad \nu = -\frac{z}{a},$$
 (8.2-9)

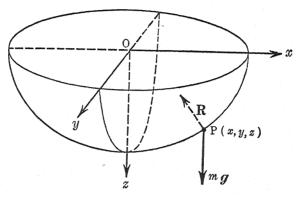


Fig. 8.2

if a is the radius of the bowl. Moreover

$$F_x = 0$$
, $F_y = 0$, $F_z = mg$. (8.2-10)

Therefore the equations of motion (8·2-3) become

$$m\ddot{x} = -\frac{Rx}{a}$$
, $m\ddot{y} = -\frac{Ry}{a}$, $m\ddot{z} = mg - \frac{Rz}{a}$. (8.2-11)

The "power" equation [i.e. $(8\cdot2-4)$ — that which gives the time rate of change of energy] then becomes

$$\frac{1}{2}m\frac{d}{dt}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) = mg\dot{z}. \tag{8.2-12}$$

Incidentally we can readily verify that

$$\dot{x}\lambda + \dot{y}\mu + \dot{z}\nu = 0.$$

a relation which we have just shown must hold from general considerations. In the present case it becomes

$$\dot{x}x + yy + \dot{z}z = 0. (8.2-13)$$

But since the equation of the surface is

$$x^2 + y^2 + z^2 = a^2, (8.2-14)$$

eq. (8·2-13) follows at once by differentiation with respect to the time. Integration of eq. (8·2-12) gives

$$\frac{1}{2}v^2 = gz + c, (8.2-15)$$

where c is a constant of integration. If the initial value of z is z_0 and the initial velocity v_0 , then $c = \frac{1}{2}v_0^2 - gz_0$. The case where $v_0 = 0$ for $z = z_0$ is a special one and will be treated later. Here we wish to discuss the general case. Now eliminating R among the equations (8·2-11) we have

$$\frac{\ddot{x}}{x} = \frac{\ddot{y}}{y} = \frac{(\ddot{z} - g)}{z}$$
 (8·2–16)

The first equation can be integrated once. Thus

$$y\dot{x} - x\dot{y} = b, \tag{8.2-17}$$

where b is an arbitrary constant. It is not easy to integrate the other equation, but we can obviate this difficulty by recalling from eq. (8.2-13) that

$$x\dot{x} + y\dot{y} = -z\dot{z}. \tag{8.2-18}$$

Now if we square the eqs. $(8\cdot2-17)$ and $(8\cdot2-18)$ and add, the result is

$$(x^2 + y^2)(\dot{x}^2 + \dot{y}^2) = b^2 + z^2 \dot{z}^2.$$
 (8.2–19)

This may be reduced very easily to the form

$$\dot{z}^2 + \dot{y}^2 + \dot{z}^2 = \frac{b^2 + a^2 \dot{z}^2}{a^2 - z^2} \cdot \tag{8.2-19a}$$

Combining $(8\cdot2-19a)$ with $(8\cdot2-15)$, however, gives

$$\dot{z}^2 = \frac{2(gz+c)(a^2-z^2)-b^2}{a^2} \cdot \tag{8.2-20}$$

On separating the variables

$$\frac{dz}{\sqrt{2(gz+c)(a^2-z^2)-b^2}} = \frac{dt}{a}.$$
 (8·2-21)

The integral involved here is an elliptic integral. Much may be learned about the resultant motion without actually evaluating it. Thus we note first that the expression under the radical must be real. Let us investigate its roots, noting that in our problem z is

necessarily restricted to the region $0 \le z \le a$. If we call the expression, $\Psi(z)$, the following relations are true

$$\Psi(a) = -b^2, \quad \Psi(-a) = -b^2.$$
 (8·2-22)

while

$$\Psi(z_0)>0,$$

and hence z may not in general take on the value a. Since $\Psi(z)$ is positive for $z=z_0$ and negative for z=a, it must have at least one root between z_0 and a. Moreover since $\Psi(-a)$ is also negative, there is another root between $z=z_0$ and z=-a. As a matter of fact, there is a third root between z=-a and $z=-\infty$, since $\Psi(-\infty)=+\infty$. Call the three roots α , β , and γ . In the type of motion in which we are interested both α and β are positive and the result is that the motion of the particle takes place between the two horizontal circles corresponding to $z=\alpha$ and $z=\beta$.

Now let

$$\frac{\alpha+\beta}{2}=f, \quad \frac{\alpha-\beta}{2}=h, \tag{8.2-23}$$

and introduce the angular variable ϕ , where

$$z = f + h\cos 2\phi. \tag{8.2-24}$$

Then

$$z - \alpha = z - (f + h) = -2h \sin^2 \phi,$$
 (8.2-25)

while

$$z - \beta = z - (f - h) = 2h \cos^2 \phi.$$
 (8.2-26)

Since

$$\Psi(z) = -2g(z-\alpha)(z-\beta)(z-\gamma).$$

it therefore follows that

$$\Psi(z) = 8gh^2 \sin^2 \phi \cos^2 \phi \cdot (z - \gamma)$$

= $8gh^2 \sin^2 \phi \cos^2 \phi \cdot (f - \gamma + h \cos 2\phi)$. (8·2-27)

We now recall from (8.2-20) that

$$\dot{z}^2 = \frac{\Psi(z)}{a^2} \cdot \tag{8.2-28}$$

But from $(8\cdot2-24)$

$$z = -2h\dot{\phi}\sin 2\phi. \tag{8.2-29}$$

The combination of (8·2-29), (8·2-28), and (8·2-27) yields finally

$$\dot{\phi}^2 = \frac{gh}{a^2k^2} (1 - k^2 \sin^2 \phi), \qquad (8.2-30)$$

where

$$k^2 = \frac{2h}{f + h - \gamma} \tag{8.2-31}$$

Separating the variables in (8·2-30) gives

$$\frac{d\phi}{\sqrt{1-k^2\sin^2\phi}} = \frac{\sqrt{gh}}{ak}dt, \tag{8.2-32}$$

so that the time corresponding to any value of ϕ is

$$t = \frac{ak}{\sqrt{gh}} \int_0^\phi \frac{d\phi}{\sqrt{1 - k^2 \sin^2 \phi}}.$$
 (8·2–33)

This expression is in the same form as $(8\cdot 1-12)$ encountered in the study of the simple pendulum. Hence as far as motion in the z direction is concerned the particle oscillates approximately as a simple pendulum. When $\phi=0$, z=f+h, while for $\phi=\pi/2$, z=f-h. Since in our case both α and β are positive, if $\alpha>\beta$ f and h are both positive and the integration from $\phi=0$ to $\phi=\pi/2$ corresponds to the motion from $z=\alpha$ to $z=\beta$, i.e., a halfperiod. We then have, if P represents the period of the motion in the z direction,

$$\frac{P}{2} = \frac{ak}{\sqrt{gh}} K, \qquad (8.2-34)$$

where

$$K = \int_0^{\pi/2} \frac{d\phi}{\sqrt{1 - k^2 \sin^2 \phi}} = F\left(k, \frac{\pi}{2}\right).$$

Now the quarter period of a simple pendulum is given (see Sec. 8·1) by

$$\frac{P}{4} = \sqrt{\frac{l}{g}}K. \tag{8.2-35}$$

Hence the length of the simple pendulum which will have the same period as the motion in the z direction of the particle here con-

sidered will be given by

$$\frac{2ak}{\sqrt{gh}} = 4\sqrt{\frac{l}{g}},$$

$$l = \frac{a^2k^2}{4b}.$$
(8.2-36)

or

A more detailed discussion together with the treatment of the motion in x and y with diagrams of actual experiments will be found in Webster's Dynamics (p. 48 ff.).

We ought, however, to note the special case where $v_0 = 0$ for $z = z_0$ [see eq. (8·2–15)]. This corresponds to the case where the

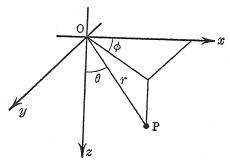


Fig. 8.3

particle is let go from *rest* at any point. What is the result? In this case we have

$$c = -gz_0. (8.2-37)$$

But since \dot{z} must also be zero initially (i.e., when $z=z_0$) it follows from eq. (8·2–20) that

$$b = 0.$$
 (8.2–38)

If now we introduce spherical coördinates defined as in Fig. 8-3, with

$$x = a \sin \theta \cos \phi,$$

 $y = a \sin \theta \sin \phi,$
 $z = a \cos \theta,$ (8.2-39)

eq. $(8\cdot2-17)$ with b=0 becomes

$$-a^2\sin^2\theta\cdot\dot{\phi}=0. \tag{8.2-40}$$

Now let us look at eq. (8.2-15) which since

$$v^2 = a^2(\dot{\theta}^2 + \sin^2\theta \cdot \dot{\phi}^2), \tag{8.2-41}$$

becomes [recalling (8·2-40)]

$$a^{2}\dot{\theta}^{2} = 2(gz + c)$$

= $2(ga\cos\theta + c)$. (8.2-42)

If now we differentiate with respect to the time we have

$$2a^2\dot{\theta}\ddot{\theta} = -2ag\dot{\theta}\sin\theta,$$

or

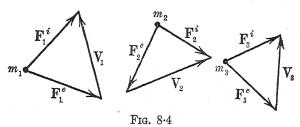
$$\ddot{\theta} + \frac{g}{a}\sin\theta = 0. \tag{8.2-43}$$

But this is the equation of the simple pendulum swinging in a plane — the problem already discussed fully in Sec. 8-1. Hence when the particle is allowed to drop from rest from its initial position the motion is that of a simple pendulum. The more complicated motion previously discussed results when the initial motion implies an initial velocity different from zero and arbitrarily directed. It should be noted, of course, that even if the initial velocity is not zero, if it is directed in the great circle containing the lowest point of the bowl, the motion is that of a simple pendulum. For in this case the particle must remain in the initial diametral plane.

8.3. Constraints and the Principles of Mechanics. D'Alembert's Principle — Dynamics Reduced to Statics. Thus far our study of dynamics has been based on our interpretation of Newton's three laws of motion outlined in Chapter I. It is important to realize, however, that there are other ways of stating the principles of mechanics. Two of these are closely connected with the motion of particles subject to constraints and hence we may well consider them at this place. The first is the celebrated principle of D'Alembert, enunciated in 1743, which provides a foundation for mechanics that effectively reduces dynamics to statics. We shall confine our attention to this principle in the present section and use Mach's method of presenting it.¹

¹Ernst Mach, "The Science of Mechanics" (English translation, Open Court Publishing Company, Chicago). An extremely valuable storehouse of information on the principles of mechanics. Every serious student of mechanics should read it.

Consider the system of n particles of masses m_1, m_2, \ldots, m_n , three of which are indicated in Fig. 8.4. Suppose that the forces $\mathbf{F}_1^i, \mathbf{F}_2^i, \ldots, \mathbf{F}_n^i$ act on them respectively We shall call these the *impressed* forces. If the masses were subject to no constraints (i.e., if they were not connected in any way or forced to move along certain curves or surfaces) they would then move with accelerations given by $\mathbf{a}_1 = \mathbf{F}_1^i/m_1$, etc. Let us suppose, however, that



due to the constraints the actual motions are such as would be produced in free bodies by the action of the forces F_1^e , F_2^e ,..., F_n^e . That is to say, $F_n^e = m_n a_n$, where a_n is the actual acceleration of the *n*th particle. These may be termed the effective forces. Now let us form the vector differences

$$F_{1}^{i} - F_{1}^{e} = V_{1},
 F_{2}^{i} - F_{2}^{e} = V_{2},
 \vdots
 \vdots
 F_{n}^{i} - F_{n}^{e} = V_{n}.$$
(8.3-1)

D'Alembert's principle then consists in the hypothesis that if the system of forces $V_1, V_2, ..., V_n$ (sometimes referred to as "lost" forces) alone were to act the system would remain in equilibrium. This can be expressed analytically by the use of the principle of virtual work (Secs. 5.7 and 5.8) in the form

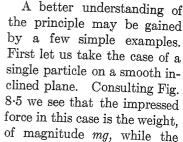
$$\sum_{j=1}^{n} V_j \cdot dr_j = 0. (8.3-2)$$

It is important to emphasize that (8.3-2) does not by any means necessarily mean that $\sum V_i = 0$. The dynamical problem is essentially reduced to one in static equilibrium. In the employment of D'Alembert's principle the essential problem is the

correct choice of the impressed forces. The constraint forces do not enter, but one must be careful not to treat a genuine impressed

force as a constraint force and leave it out.

A better understanding of



T.

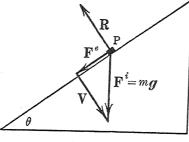


Fig. 8.5

effective force is F^e acting along the plane. In this case there is but one particle and a single $V = F^i - F^e$.

Application of (8·3-2) leads at once to

$$(\mathbf{F}^i - \mathbf{F}^e) \cdot d\mathbf{r} = 0, \qquad (8.3-3)$$

where dr is taken down the plane.

Hence

$$\mathbf{F}^i \cdot d\mathbf{r} = mg \sin\theta \, dr,$$

while

or

$$\mathbf{F}^e \cdot d\mathbf{r} = F^e dr = ma dr.$$

We therefore have finally since dr is arbitrary,

$$F^e = ma = mg \sin \theta,$$
 m_1 $a = g \sin \theta,$ (8.3-4) Fig. 8.6

the usual expression for the acceleration of the particle down the plane. This is doubtless a rather long-winded way of getting at a simple result. Nevertheless it is often desirable to emphasize the meaning of a general principle by the use of a very simple example. Let us try another illustration, this time one involving two particles. The simplest case of this kind is probably Atwood's machine, already worked out in Sec. 2·1 by means of Newton's laws. Consulting Fig. 8·6 we see that the impressed forces are the weights $m_1 \mathbf{g}$ and $m_2 \mathbf{g}$ respectively. The tensions \mathbf{T}_1 and \mathbf{T}_2 are constraint

forces and do not enter the expression of the principle. The latter now takes the form

$$(m_1\mathbf{g} - m_1\mathbf{a}_1) \cdot d\mathbf{r}_1 + (m_2\mathbf{g} - m_2\mathbf{a}_2) \cdot d\mathbf{r}_2 = 0.$$
 (8.3–5)

If we follow the pattern of Sec. 5.8 we find it convenient to represent $d\mathbf{r}_1$ and $d\mathbf{r}_2$ as both directed vertically. The expansion of the dot product in (8.3-5) then yields

$$(m_1g - m_1a_1) dr_1 + (m_2g - m_2a_2) dr_2 = 0.$$
 (8.3-6)

But from the geometry of the constraints we must have

$$dr_1 = -dr_2, \quad a_1 = -a_2.$$
 (8.3–7)

Therefore the solution (with dr_1 arbitrary) follows in the usual form

$$a_1 = \frac{m_1 - m_2}{m_1 + m_2} g. (8.3-8)$$

The reader should work out other and more complicated examples using the principle. (See, for instance, the problems at the end of the chapter.)

In advanced texts on mechanics D'Alembert's principle is often made the starting point for the development of the whole subject. This is true, for example, in the celebrated treatise "Mechanique Analytique" of Lagrange (1811). This work should be of considerable historical interest to the student, for it marks the climax of the eighteenth century attempts to make mechanics a branch of mathematical analysis. In the preface the author proudly boasts: "There are no figures in this work." It constitutes a monument of analysis. We shall have occasion in Chapter XII to refer to Lagrange's form of the equations of motion of a dynamical system.

8.4. Gauss' Principle of Least Constraint. Another very significant principle of mechanics closely connected with the idea of motion subject to constraint is due to Gauss. It was elaborated by him in 1829 in an attempt to reduce all mechanics to a single generalization.

Suppose that we have given a system of n particles of masses $m_1, m_2, m_3, \ldots, m_n$ occupying the positions A_1, \ldots, A_n . Three

of them are indicated in the accompanying figure (Fig. 8.7). Let us now imagine that if the particles were perfectly free to move under the action of certain external forces, they would undergo during an infinitesimal time interval $d\tau$ the infinitesimal displacements $\overline{A_1B_1}$, $\overline{A_2B_2}$, ..., $\overline{A_nB_n}$. The system being subjected to certain constraints (i.e., the masses being perhaps connected to

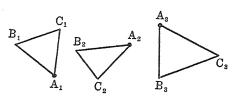


Fig. 8.7

each other by rods or cords, or constrained to move along certain curves or surfaces), suppose that the actual displacements during the above time interval are $\overline{A_1C_1}$, $\overline{A_2C_2}$, ..., $\overline{A_nC_n}$. Gauss' principle now states that the actual displacement of the system under the constraints is such that the sum

$$m_1\overline{B_1C_1}^2 + \cdots + m_n\overline{B_nC_n}^2,$$
 (8.4-1)

is a minimum, that is, less for the actual motion than for any other possible motion of the same system under the same constraints. This sum may be looked upon as the analytical expression for the total constraint, and hence the principle is known as that of least constraint. It is interesting to note that the static case is included in this principle, for if the total constraint for every possible motion is greater than it would be for rest, equilibrium will prevail, i.e., the system once at rest will remain at rest.

Gauss intended this as a fundamental postulate by the assumption of which all the problems of mechanics can be solved. It is necessarily connected with the Newtonian laws of motion and D'Alembert's principle, and indeed we shall show that it may be deduced from the latter. Nevertheless for the moment let us assume it and work out a fairly simple problem in dynamics to understand its essential content.

We choose the Atwood machine again. Examining the situation once more as in Fig. 8.8 we note that if m_1 and m_2 were free of

constraint, in the time $d\tau$ they would each fall vertically from rest the distance

$$s = \frac{1}{2}g(d\tau)^2. \tag{8.4-2}$$

Actually, however, m_1 falls with acceleration of magnitude a and hence in time $d\tau$ travels the distance

$$s' = \frac{1}{2}a(d\tau)^2. \tag{8.4-3}$$

The total constraint is thus

$$\frac{m_1}{4} (a-g)^2 (d\tau)^4 + \frac{m_2}{4} (-a-g)^2 (d\tau)^4.$$
(8·4-4)

This, according to the principle, is to be a minimum. Hence, differentiating with respect to a and setting the result equal to zero, we have

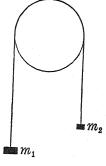


Fig. 8.8

$$\frac{m_1}{2} (a-g)(d\tau)^4 + \frac{m_2}{2} (a+g)(d\tau)^4 = 0,$$

or

$$a(m_1 + m_2) = g(m_1 - m_2),$$
 (8.4-5)

and therefore

$$a = \frac{m_1 - m_2}{m_1 + m_2} g, (8.4-6)$$

as usual.

Let us now show that the principle of least constraint follows from that of D'Alembert.

Referring again to Fig. 8·7, we recall that $\overline{A_iB_i}$ is the displacement which the mass m_j would undergo under the application of the impressed force $F_j{}^i$ if it were free to do so. $\overline{A_jC_j}$ is the displacement it actually undergoes under the action of the effective force $F_j{}^e$, the deviation due to the constraints being $\overline{B_jC_j}$. Now imagine (see Fig. 8·9) that instead of going to C_j the particle were to go to D_j with a corresponding deviation $\overline{B_jD_j}$. Let us denote $\overline{C_jD_j}$ by δr_j , and suppose it makes the angle θ_j with $\overline{C_jB_j}$. We then have

$$\overline{B_i D_i^2} = \overline{C_i B_i^2} + \delta r_i^2 - 2 \, \delta r_i \overline{C_i B_i} \cos \theta_i, \qquad (8.4-7)$$

whence multiplying by m_j and summing up over all the particles there results

$$\sum m_{j} \overline{B_{j}} \overline{D_{j}}^{2} - \sum m_{j} \overline{C_{j}} \overline{B_{j}}^{2} = \sum m_{j} \delta r_{j}^{2}$$

$$- 2 \sum m_{j} \overline{C_{j}} \overline{B_{j}} \delta r_{j} \cos \theta_{j}.$$
(8.4–8)

Now since all the above displacements are supposed to take place from rest in the same time $d\tau$, it follows that

$$\overline{A_j B_j} = k a_{ij}, \quad \overline{A_j C_j} = k a_{ej}, \quad \overline{C_j B_j} = k a_{ej}, \quad (8.4-9)$$

where k is a constant and a_{ij} and a_{ej} are the magnitudes of the accelerations produced by the impressed and effective forces

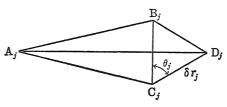


Fig. 8.9

respectively and a_{ci} is the magnitude of the acceleration that would correspond to the particular constraint which acts here (assuming it were able to produce an acceleration). Thus we may think of the actual (effective) acceleration as produced by subtracting an "equivalent" acceleration due to the constraint from the acceleration that would be produced in the particle if it were free to move. Then

$$m_j \overline{A_j B_j} = k F_j^i, \quad m_j \overline{A_j C_j} = k F_j^e, \quad m_j \overline{C_j B_j} = k V_j. \quad (8.4-10)$$

Now according to D'Alembert's principle

$$\sum \mathbf{V}_j \cdot \delta \mathbf{r}_j = 0. \tag{8.4-11}$$

Therefore from eq. (8.4-10)

$$2\sum m_i \overline{C_i B_i} \, \delta r_i \cos \theta_i = 0, \qquad (8.4-12)$$

and finally from (8.4-8)

$$\sum m_j \overline{B_j D_j}^2 - \sum m_j \overline{C_j B_j}^2 = \sum m_j \, \delta r_j^2. \tag{8.4-13}$$

The right-hand side of this equation is always positive.

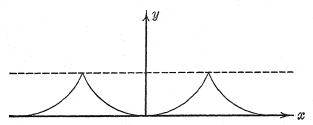
It follows then from (8.4-13) that the $\sum m_j \overline{C_j} B_j^2$ must be less for the actual motion than for any other alternative motion

compatible with the constraints. This is the essential content of the principle of least constraint.

The student will do well to apply the principles of D'Alembert and least constraint to other simple examples. In a later portion of the book it will be shown how still more general mechanical principles may replace or supplement them. Their introduction in this place is for the purpose of emphasizing the important fact that there are several ways of expressing the laws of mechanics, depending on the point of view which is taken.

PROBLEMS

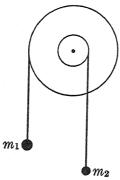
- 1. A particle is constrained to move under gravity in a vertical circle of radius R. If it starts from rest where the tangent to the circle is vertical, how long will it take it to reach the lowest point of the circle?
- 2. In Problem 1, find the expression for the reaction of the circle on the particle as a function of vertical distance above the lowest point.
- 3. A particle of mass m is fixed to one end of a string of length l and the other end is attached to a fixed point. The particle is then constrained to move in a horizontal circle with uniform angular velocity ω . Find the inclination of the string to the vertical and the relation between the tension in the string and the angular velocity. Compute the length of the equivalent simple pendulum. (This arrangement is known as a conical pendulum. Suggest a practical use of it.)
- 4. The cycloid is the locus of a fixed point on a circle which rolls along a fixed axis. Show that the parametric equations of an inverted cycloid with its lowest point at the origin (see figure) are $x = a(\theta + \sin \theta)$, $y = a(1 \cos \theta)$. Find the time it takes for a particle to descend under gravity from rest from



any point on an *inverted* cycloid to the lowest point. Show that this time is independent of the point from which the descent takes place. Comment on the significance of this result and its possible practical application.

5. A particle moves under the influence of gravity on the convex side of a vertical circle with center at the ground. If it starts from rest at the top, where will it leave the circle?

6. A particle is constrained to move in a straight line under the influence of a force directed toward a fixed point at a perpendicular distance a from the line and varying inversely as the square of the distance from the point. Discuss the motion and in particular compute the approximate time required by the particle to travel from x = h to x = 0, where x = 0 corresponds to the position of the fixed point, and $h \ll a$.



- 7. A smooth inclined plane of mass M and angle α is free to move on a perfectly smooth horizontal plane. A particle of mass m is free to move on the inclined plane. Discuss the motion of both particle and inclined plane.
- 8. Use the principle of D'Alembert to obtain the motion of the simple pendulum. Solve the same problem with the principle of least constraint.
- 9. Solve the problem of the simple lever by the principles of D'Alembert and least constraint.
- 10. A mass m_1 is attached to a string wrapped about a wheel of mass M and radius R, while a mass m_2 is attached to a string wrapped about

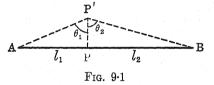
the axle of mass m and radius r. Find the accelerations of m_1 and m_2 respectively by the principles of D'Alembert and least constraint.

CHAPTER IX

OSCILLATIONS

9.1. A Simple Problem in Vibration. Some of the most important cases of constrained motion are those in which the particles are connected by rods and strings. Thus consider the following simple example later to be generalized. Imagine (Fig. 9.1) a particle of mass m attached at the point P to a horizontal, elastic string (whose mass is so small in comparison with m that it may be considered negligible) fastened at the two ends A and B and stretched with tension τ .

The distances AP and PB are represented by l_1 and l_2 respectively. The particle is pulled aside a *short* distance PP' perpendicular to AB and let go. We are to find the



resulting motion, assumed to be rectilinear. Call the displacement at any instant ξ and suppose it is very much less than $l(=l_1+l_2)$. We shall make the simplifying assumption that the alteration in the tension with the stretching of the string is too slight to be taken into account. The total vertical component of force on the particle is then

$$-\tau(\cos\theta_1+\cos\theta_2). \tag{9.1-1}$$
 Now

 $\cos \theta_1 = \frac{\xi}{\sqrt{\xi^2 + \overline{l_1}^2}},$ $\cos \theta_2 = \frac{\xi}{\sqrt{\xi^2 + \overline{l_2}^2}}.$ (9.1-2)

But if $\xi \ll l_1$ and $\xi \ll l_2$, as we have assumed, we can write to a sufficiently good approximation

$$\cos \theta = \frac{\xi}{l_1}, \quad \cos \theta_2 = \frac{\xi}{l_2}, \tag{9.1-3}$$

so that the equation of motion of the particle becomes

$$m\ddot{\xi} = -\tau \xi \left(\frac{1}{l_1} + \frac{1}{l_2}\right) \cdot \tag{9.1-4}$$

This is the equation of simple harmonic motion (see Sec. 2.2) with frequency

$$\nu = \frac{1}{2\pi} \sqrt{\frac{\tau(l_1 + l_2)}{ml_1 l_2}} = \frac{1}{2\pi} \sqrt{\frac{\tau l}{ml_1 l_2}},$$
 (9·1–5)

where l is the total length of the string. If $l_1 = l_2 = l/2$, we get the simple case

$$\nu = \frac{1}{\pi} \sqrt{\frac{\tau}{ml}}.$$
 (9·1-6)

This problem should be compared with the simple pendulum and the simple spring treated previously (Secs. 2·2 and 8·1). It affords a simple illustration of the small oscillations of a dynamical system about a position of equilibrium. We may use it indeed as an introduction to a more extended discussion of this important topic.

9.2. Oscillations of a Dynamical System with One Degree of Freedom — Dissipation. We have had occasion in several preceding sections so far to note the occurrence of simple harmonic oscillations of small amplitude. We need recall only the simple pendulum, the simple spring and the mass particle fixed on a horizontal string (Sec. 9.1). Such oscillations are so important throughout all physics that it will be desirable to discuss their general features with greater elaboration.

We shall at first restrict ourselves to the case of a system with a single degree of freedom. By this is meant a particle or collection of particles whose state of motion at any instant is completely determined by one variable, which may of course be quite general in nature, i.e., a linear displacement, angular displacement or some complicated combination of these. In our present discussion we shall be content to use a linear displacement as our variable, and shall denote it by the letter ξ . Associated with the system will be a certain mass designated by m and a certain stiffness denoted by f. The meaning of the latter is as follows. Sup-

pose that a static force F_1 is necessary to produce a static displacement of the system of magnitude ξ_1 . Then

$$f = \frac{F_1}{\xi_1} \tag{9.2-1}$$

is defined as the *stiffness* coefficient or more simply the stiffness of the system.¹ If now the latter is displaced a slight amount from its equilibrium position and then allowed to move freely the equation of motion is clearly

$$m\ddot{\xi} = -f\xi, \tag{9.2-2}$$

leading to simple harmonic motion with frequency $\nu = \frac{1}{2\pi} \sqrt{\frac{f}{m}}$

and period $P = 2\pi \sqrt{\frac{m}{f}}$. Thus in the case of the vibration prob-

lem treated in Sec. 9·1, the effective stiffness is represented by $f = \tau l/l_1 l_2$, while the mass is the mass of the particle. It may be pointed out that there are certain cases where the equivalent mass (i.e., that entering into the formula) is *not* the whole mass of the oscillating system. This is notably true in the case of a membrane or diaphragm.

However, at present we are more interested in another matter. We have already noted the fact that the eq. (9·2-2) can never adequately represent the true state of affairs for any oscillatory motion occurring in nature, since it implies that once the system is oscillating it continues so indefinitely; this is contrary to experience. We must recognize that there is always present a certain resistance to the motion leading to dissipation, so that ultimately unless some external influence intervenes, rest results. The question arises: how shall this resistance be introduced into the equation of motion? Obviously the resistance may be thought of as equivalent to a force, in nature somewhat like the force of friction in that it is always opposite in direction to the motion. It can hardly be supposed to be proportional to the displacement or any function thereof. On the other hand it does seem reasonable

¹ It is here assumed that the stiffness remains constant, i.e., every time the static force is F_1 the corresponding displacement is ξ_1 . This implies, as we shall see later in Chap. X, that the system is elastic. But see the comments at the end of this section.

to assume that it is a function of the velocity. Let us therefore suppose that in addition to the restoring force of magnitude $-f\xi$ due to the stiffness there is present a resistance force of magnitude $-R\xi$, where R will be termed the damping factor or coefficient. It must be emphasized that this is a pure hypothesis, since we have no a priori justification for the first power rather than the second or higher. And the essential confirmation of the correctness of this assumption will appear only after the equation of motion thus constructed has been integrated and the resulting motion compared with experience.

The equation of motion on this hypothesis now becomes

$$m\ddot{\xi} = -f\xi - R\dot{\xi},$$

or more simply

$$m\ddot{\xi} + R\dot{\xi} + f\xi = 0.$$
 (9.2–3)

The solution of this equation is given in texts on differential equations, but we may proceed rather simply as follows. The form of the left-hand side indicates an exponential function of the time. Hence we assume for our solution

$$\xi = Ae^{\lambda t}, (9.2-4)$$

where A is an arbitrary constant and λ is to be determined. We then have

$$\dot{\xi} = A\lambda e^{\lambda t}, \quad \ddot{\xi} = A\lambda^2 e^{\lambda t},$$

and therefore on substitution

$$m\lambda^2 + R\lambda + f = 0. (9.2-5)$$

This is the condition which λ must satisfy in order that (9·2-4) may be a solution of the equation (9·2-3). Solving we find

$$\lambda = -\frac{R}{2m} \pm \sqrt{\frac{R^2}{4m^2} - \frac{f}{m}},$$
 (9.2-6)

and consequently the solution becomes

$$\xi = e^{-R/2m \cdot t} (A e^{\sqrt{R^2/4m^2 - f/m} \cdot t} + B e^{-\sqrt{R^2/4m^2 - f/m} \cdot t}), \quad (9 \cdot 2 - 7)$$

where now the *two* arbitrary constants A and B must enter since we get a solution with the minus sign as well as with the plus sign.

There are three special cases to consider according as

$$\frac{R^2}{4m^2} \ge \frac{f}{m} \,. \tag{9.2-8}$$

Let us take first the case where $R^2/4m^2 < f/m$, i.e., where the damping factor R is relatively small. The radical is then imaginary and we may write

$$\xi = e^{-R/2m \cdot t} (Ae^{i\sqrt{f/m - R^2/4m^2} \cdot t} + Be^{-i\sqrt{f/m - R^2/4m^2} \cdot t}), \quad (9 \cdot 2 - 9)$$

with $i = \sqrt{-1}$. It will be necessary here to recall one or two facts about complex numbers. Any such number may be represented by a point P in the xy plane with coördinates x and y. The number itself is then written

$$z \equiv x + iy. \qquad (9.2-10)$$

But there is another way of expressing this, namely in terms of polar coördinates, for $x = r \cos \theta$ and $y = r \sin \theta$ (Fig. 9.2), whence

$$r$$
 θ
 x

$$z = r(\cos\theta + i\sin\theta)$$
. (9·2–11)

The angle θ is called the *amplitude* or *argument* of the complex number while r is called its *modulus* or *absolute value*. Thus

$$r = \sqrt{x^2 + y^2}.$$

Now consider the complex number

$$w = \cos \theta + i \sin \theta. \tag{9.2-12}$$

This has the modulus unity and we may represent it by a point P in our diagram for which r=1. If we displace P to P' on the unit circle, w changes to w+dw and θ changes to $\theta+d\theta$. But we have from $(9\cdot 2-12)$ at once

$$dw = (-\sin\theta + i\cos\theta) d\theta$$
$$= i(\cos\theta + i\sin\theta) d\theta, \qquad (9.2-13)$$

whence

$$dw = iw \, d\theta. \tag{9.2-14}$$

Integrating,

$$\log w = i\theta + C,$$

or

$$w = C'e^{i\theta}, (9.2-15)$$

and since w = 1 for $\theta = 0$, C' = 1 and hence

$$w = e^{i\theta}. (9.2-16)$$

This extremely important connection between the imaginary exponential and the circular functions enables us to write at once in place of (9.2-9)

$$\xi = e^{-R/2m \cdot t} [(A + B) \cos \gamma t + (A - B)i \sin \gamma t], \quad (9.2-17)$$

where for simplicity we have set $\gamma = \sqrt{f/m - R^2/4m^2}$. In expressing the result in this form it is necessary to admit that A and B may be complex numbers. Write $A = a_1 + ia_2$ while $B = b_1 + ib_2$. Then if we separate ξ into real and imaginary parts we have

$$\xi = e^{-R/2m \cdot t} [(a_1 + b_1) \cos \gamma t - (a_2 - b_2) \sin \gamma t + i\{(a_2 + b_2) \cos \gamma t + (a_1 - b_1) \sin \gamma t\}]. \quad (9.2-18)$$

From the form of the solution it is seen that both the real and imaginary parts are equally solutions and moreover are essentially of the same form. We shall choose to use the real part (and indeed make this our general rule in problems of this kind), and writing $a_1 + b_1 = A'$ and $a_2 - b_2 = B'$, our result is

$$\xi = e^{-R/2m \cdot t} [A' \cos \gamma t - B' \sin \gamma t]. \tag{9.2-19}$$

This result can be put into even more compact form if we introduce the new variables C and ϵ where

$$A' = C \cos \epsilon$$
 and $B' = C \sin \epsilon$. (9.2-20)

With these new constants we get for the displacement of the system

$$\xi = Ce^{-R/2m \cdot t} \cos (\gamma t + \epsilon). \tag{9.2-21}$$

We are now ready to interpret the final answer physically. If we plot ξ as a function of t the result is indicated in Fig. 9.3 and represents an oscillation of frequency

$$\nu = \frac{\gamma}{2\pi},\tag{9.2-22}$$

265

the amplitude of which decays steadily with the progress of time. It is often called a damped oscillation. Two arbitrary constants are involved: C, the initial amplitude, and ϵ , the initial phase. By definition the phase of the motion at any instant t is

$$\phi = \gamma t + \epsilon. \tag{9.2-23}$$

With regard to the frequency we note that if the damping is small, i.e., if $\bar{R}^2/4m^2 \ll f/m$, the frequency does not differ noticeably

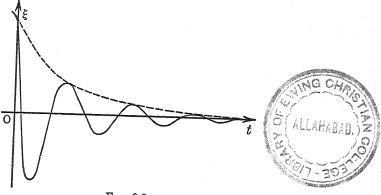


Fig. 9.3

from that of the corresponding undamped oscillator. Thus slight damping has little effect on the frequency of a simple harmonic This is well borne out by experiments on pendulums, springs, tuning forks, etc., and serves as one check on the reasonableness of our equation of motion for the oscillator.

The magnitude of the damping effect is given by the term R/2m. The reciprocal of this quantity, viz., 2m/R, is often called the "decay modulus": it represents the time taken for the oscillation amplitude to decay to 1/eth of its original value. damping it is accordingly very large. Another commonly used measure of the damping is the so-called "logarithmic decrement." It is the logarithm to the base e of the ratio of two successive amplitudes. Thus if $P(=1/\nu)$ is one period of the motion,

$$\log \frac{Ce^{-R/2m \cdot t_1}}{Ce^{-R/2m \cdot (t_1 + P)}} = \frac{R}{2m} P$$
 (9·2-24)

is the logarithmic decrement.

It will be worth while to pause and consider a little the evalua-

tion of the constants C and ϵ from the initial conditions. Suppose that when t = 0, $\xi = \xi_0$. We then have from (9.2-21)

$$\xi_0 = C \cos \epsilon. \tag{9.2-25}$$

This alone is not sufficient to determine both constants. We might of course note the displacement at some later time and so obtain another relation from which together with $(9\cdot2-25)$ C and ϵ may be calculated. Actually it is simpler to use the initial ve-locity for this purpose. Thus differentiating $(9\cdot2-21)$ with respect to the time yields

$$\dot{\xi} = C\left(-\frac{R}{2m}\right)e^{-R/2m\cdot t}\cos\left(\gamma t + \epsilon\right) - Ce^{-R/2m\cdot t}\gamma\sin\left(\gamma t + \epsilon\right). \quad (9\cdot2-26)$$

Now when t = 0, if $\dot{\xi} = \dot{\xi}_0$ we have

$$\dot{\xi}_0 = -\frac{RC}{2m}\cos\epsilon - C\gamma\sin\epsilon, \qquad (9.2-27)$$

whence we find using (9.2-25)

$$\tan \epsilon = -\frac{\dot{\xi}_0 + \frac{R}{2m} \, \xi_0}{\gamma \xi_0}, \qquad (9.2-28)$$

while C is finally obtained as

$$C = \frac{\sqrt{\gamma^2 \xi_0^2 + \left(\dot{\xi}_0 + \frac{R}{2m} \xi_0\right)^2}}{\gamma}.$$
 (9.2-29)

In case the damping is so slight as to be negligible, i.e., $R/2m \ll 1$, we have approximately

$$\tan \epsilon = -\frac{\dot{\xi}_0}{\xi_0 \sqrt{\frac{f}{m}}}, \qquad (9.2-30)$$

and

$$C = \sqrt{\frac{\xi_0^2 \frac{f}{m} + \xi_0^2}{\frac{f}{m}}}$$

$$= \sqrt{\xi_0^2 + \frac{m}{f} \xi_0^2}.$$
(9.2-31)

Before entering upon further discussion of these damped oscillations and their interesting analogies throughout the realm of physics we must dispose of the other two cases indicated in (9·2–8). Take first the rather unusual case where

$$\frac{f}{m} = \frac{R^2}{4m^2} \cdot \tag{9.2-32}$$

Here the roots of $(9\cdot2-5)$ are both equal to -R/2m and our procedure must be somewhat modified. Instead of substituting $\xi = Ae^{\lambda t}$ as in $(9\cdot2-4)$ we let

$$\xi = e^{\lambda t} \chi(t), \qquad (9.2-33)$$

where $\chi(t)$ is at first an undetermined function of t. Then

$$\dot{\xi} = \lambda e^{\lambda t} \chi + e^{\lambda t} \dot{\chi},
\dot{\xi} = \lambda^2 e^{\lambda t} \chi + 2\lambda e^{\lambda t} \dot{\chi} + e^{\lambda t} \ddot{\chi}.$$

and on substitution into the equation of motion

$$m(\lambda^2 \chi + 2\lambda \dot{\chi} + \ddot{\chi}) + R(\lambda \chi + \dot{\chi}) + f\chi = 0, \quad (9.2-34)$$

or rearranging

$$\chi(m\lambda^2 + R\lambda + f) + \dot{\chi}(2m\lambda + R) + m\ddot{\chi} = 0. \quad (9.2-35)$$

Now since (9·1-5) is satisfied by $\lambda = -R/2m$, the above reduces to

$$\ddot{\chi} = 0.$$
 (9.2–36)

Therefore

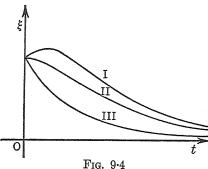
$$\chi = A + Bt, \qquad (9.2-37)$$

where A and B are arbitrary constants, will satisfy the equation. The solution in this case then becomes

$$\xi = e^{-R/2m \cdot t} (A + Bt).$$
 (9.2–38)

We note at once that $(9\cdot1-31)$ does not represent an oscillation. Depending on the initial conditions there are various possible types of curves. Taking the case where A and B are both positive if A < 2mB/R, a curve of type I with a maximum displaced from t = 0 results, while if A = 2mB/R the curve is of type II with maximum at the origin, and finally if A > 2mB/R, there is no maximum but ξ decreases steadily to zero as time progresses. All

three represent what is called critically damped motion. viously if A and B are of different signs, ξ will become zero for



finite t, and somewhat more complicated curves than those shown in Fig. 9.4 will result. The reader should investigate these for himself.

We finally consider the case where $R^2/4m^2 > f/m$, i.e., γ imaginary. Here again, as inspection of eq. (9.2-7) shows, no oscillations are present. We can

represent the result somewhat more simply if we introduce some simple initial conditions, say $\xi = \xi_0$ and $\dot{\xi} = 0$ for t = 0. with $\gamma' = i\gamma$ (where γ' is real) the solution (9·2–7) is

$$\xi = e^{-R/2m \cdot t} (Ae^{\gamma' t} + Be^{-\gamma' t}).$$

The initial condition gives

$$\xi_0 = A + B,$$
 (9.2–39)

while since

$$\dot{\xi} = -\frac{R}{2m} e^{-R/2m \cdot t} (A e^{\gamma' t} + B e^{-\gamma' t})$$
$$+ e^{-R/2m \cdot t} (A \gamma' e^{\gamma' t} - B \gamma' e^{-\gamma' t}),$$

for t = 0, we have

$$0 = -\frac{R}{2m} (A + B) + \gamma'(A - B)$$
$$= A \left[\gamma' - \frac{R}{2m} \right] - B \left[\gamma' + \frac{R}{2m} \right]. \qquad (9.2-40)$$

Solving for A and B from (9.2-39) and (9.2-40), we have

$$A = \frac{\xi_0 \left[\gamma' + \frac{R}{2m} \right]}{2\gamma'},$$

$$B = \frac{\xi_0 \left[\gamma' - \frac{R}{2m} \right]}{2\gamma'}.$$

$$(9.2-41)$$

Then for the displacement there results

$$\xi = \frac{\xi_0 e^{-R/2m \cdot t}}{2\gamma'} \left[\gamma' (e^{\gamma' t} + e^{-\gamma' t}) + \frac{R}{2m} (e^{\gamma' t} - e^{-\gamma' t}) \right] \cdot (9 \cdot 2 - 42)$$

Now using hyperbolic functions we may write

$$e^{\gamma't} + e^{-\gamma't} = 2 \cosh \gamma't,$$

$$e^{\gamma't} - e^{-\gamma't} = 2 \sinh \gamma't.$$

whence it follows that the displacement is

$$\xi = \frac{\xi_0 e^{-R/2m \cdot t}}{\gamma'} \left[\gamma' \cosh \gamma' t + \frac{R}{2m} \sinh \gamma' t \right] \cdot (9 \cdot 2 - 43)$$

The reader should plot this as an exercise.

Physically the last two cases, in which $R^2/4m^2 \ge f/m$, correspond to rather considerable damping such as may be made manifest by the motion of a pendulum in a very viscous fluid like molasses or heavy tar. Strictly speaking they thus lie outside the realm of oscillations which are the theme of this chapter. Nevertheless there are some interesting analogies in other branches of physics.

A rather interesting case of damped motion occurs when the mass of the system is so small that the term $m\ddot{\xi}$ in (9·2–3) can be neglected. The equation of "motion" then becomes

$$R\dot{\xi} + f\xi = 0, \qquad (9\cdot 2-44)$$

with solution

$$\xi = \xi_0 e^{-ft/R}. \tag{9.2-45}$$

This means physically that the system has been started off with initial displacement ξ_0 and initial velocity $-\xi_0 f/R$. Obviously the displacement goes asymptotically to zero as t increases and there is no oscillation. In time $\tau = R/f$, ξ becomes ξ_0/e and it is said that the system has relaxed to 1/eth of the initial displacement. The time R/f is then called the "relaxation time."

The notion of relaxation time is important in many physical phenomena. Thus when we try to stretch a wire by hanging a weight on it the increase in length given by Hooke's law does not immediately take place. The elementary physics relation

$$F = f\xi \tag{9.2-46}$$

for the stretching force corresponding to stretch ξ assumes an equilibrium condition which does not prevail at the instant that

the force is applied. Hence to study the problem of stretching a wire dynamically we must replace (9·2–46) by an equation which takes account of the velocity with which the stretching process proceeds. The simplest form of such an equation is

$$F = f\xi + R\xi, \tag{9.2-47}$$

which for constant applied force F has the solution

$$\xi = F/f \cdot (1 - e^{-ft/R}).$$
 (9.2–48)

After time $\tau=R/f$, ξ reaches to within 1/e of F/f, which is the final stretch. Hence once more we term R/f the relaxation time of the process.

Every physical phenomenon in which there is a time lag in the attainment of the effect of a given cause will show a relaxation time. Hence the notion is of considerable importance in physics.

9.3. Energy of Damped Oscillations. Time Averages. Let us revert to the oscillations discussed in the previous section and consider them from the standpoint of the energy involved. It will be recalled that eq. $(9\cdot2-2)$ refers to the motion of a conservative system (Sec. 4·1), for on multiplying through both sides by ξdt , we have

$$d(\frac{1}{2}m\dot{\xi}^2) + d(\frac{1}{2}f\xi^2) = 0,$$

or

$$\frac{1}{2}m\dot{\xi}^2 + \frac{1}{2}f\xi^2 = C. \tag{9.3-1}$$

The first term on the left is the kinetic energy K of the oscillator, while the second term is the potential energy V, for by definition it is the work done while the oscillator is displaced an amount ξ against the restoring force $-f\xi$. The sum of the kinetic and potential energies is the *total* energy, which in this case is constant.

The situation is different, however, as soon as dissipation is taken into account. Thus if we multiply eq. (9.2-3) by $\dot{\xi}$, we obtain

$$\dot{K} + \dot{V} = -R\dot{\xi}^2 = \dot{U},\tag{9.3-2}$$

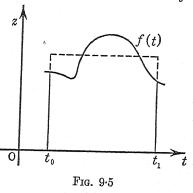
where the total energy is represented by U. In other words the energy of the system is decreasing with the time at the rate $R\xi^2$. This is then the rate of dissipation of energy and one-half of it has been called by Lord Rayleigh¹ the dissipation function. The

^{1 &}quot;Theory of Sound," Vol. I, §81.

system under consideration is no longer a conservative but a dissipative system.

Let us digress for a moment to note that an important concept in connection with oscillating systems is that of time average.

Just as one forms the arithmetical mean or average of a number of similar quantities by adding them and dividing by the number, so one may form the average of the function of a variable over a given range or interval of values of the variable by integrating the function over this interval and dividing the result by the magnitude of the interval. Considering the matter geo-



metrically, if z = f(t) is the function and we plot it as in the attached illustrative figure (Fig. 9.5), the average of f(t) over the

interval from $t = t_0$ to $t = t_1$ is denoted by $\overline{f(t)}$ (the t above the bar may be left off if it is perfectly definite that the average is with respect to t), where

$$\overline{f(t)} = \frac{1}{t_1 - t_0} \cdot \int_{t_0}^{t_1} f(t) dt.$$
 (9.3–3)

It is thus evident that $\overline{f(t)}$ is the ordinate which when multiplied by the magnitude of the interval gives an area equal to the actual area between the curve and the t axis from t_0 to t_1 .

We may now discuss \overline{K} and \overline{V} for the case of an oscillator. Taking first for simplicity the case of an *undamped* oscillator we have

$$\xi = A \cos (2\pi \nu t + \epsilon), \tag{9.3-4}$$

where ν is the frequency, A the amplitude and ϵ the initial phase. Then

 $\xi = -2\pi\nu A \sin (2\pi\nu t + \epsilon). \tag{9.3-5}$

The kinetic and potential energies become respectively

$$K = 2\pi^2 v^2 A^2 m \sin^2(2\pi v t + \epsilon), \qquad (9.3-6)$$

$$V = \frac{1}{2}fA^2 \cos^2(2\pi\nu t + \epsilon), \tag{9.3-7}$$

and we note that K+V= constant, as must necessarily be the case. The averages over a single period $P=1/\nu$ are then

$$\bar{K} = \frac{2\pi^2 \nu^2 m A^2}{P} \int_0^P \sin^2(2\pi\nu t + \epsilon) dt, \qquad (9.3-8)$$

$$\bar{V} = \frac{fA^2}{2P} \int_0^P \cos^2(2\pi\nu t + \epsilon) dt.$$
 (9.3-9)

Now

$$\frac{1}{P}\int_0^P \sin^2\left(2\pi\nu t + \epsilon\right) dt = \frac{1}{4\pi\nu P} \left[\left(2\pi\nu t + \epsilon\right) - \frac{1}{2}\sin\left(2\pi\nu t + \epsilon\right) \right]_0^P.$$

Since $\sin 2(2\pi + \epsilon) = \sin 2\epsilon$, the above reduces to $\frac{1}{2}$, and hence $\overline{K} = \pi^2 \nu^2 m A^2$. (9.3–10)

Similarly

$$\frac{1}{P} \int_0^P \cos^2(2\pi\nu t + \epsilon) dt = \frac{1}{4\pi\nu P} \left[(2\pi\nu t + \epsilon) + \frac{1}{2} \sin 2(2\pi\nu t + \epsilon) \right]_0^P$$
$$= \frac{1}{2},$$

and therefore

$$\bar{V} = \frac{fA^2}{4} \cdot \tag{9.3-11}$$

But the frequency of the oscillation is

$$\nu = \frac{1}{2\pi} \sqrt{\frac{f}{m}},$$

so that $f = 4\pi^2 v^2 m$. It thus develops that

$$\overline{K} = \overline{V} \tag{9.3-12}$$

over a single period. Obviously the same result will hold over any number of periods. In fact it may be noted that since

$$\frac{1}{\tau} \int_0^{\tau} \frac{\sin^2}{\cos^2} \left[(2\pi\nu t + \epsilon) dt \doteq \frac{1}{2}, \qquad (9.3-13) \right]$$

as the time interval τ grows greater, even if it is not a multiple of one period, the average kinetic energy over *any* interval of time large compared with one period is equal to the average potential energy. In general in what follows, when we speak of a time

average for an oscillating system we shall mean that taken over such an interval.

The reader may find it of interest to compare with the above the *space* averages of V and K, though the latter are not usually of such physical interest as the time average.

9.4. Forced Oscillations of a Dissipative System. The oscillations discussed in the previous two sections are what may be called "free" or "natural" oscillations. Suppose now that a periodic force of frequency $\nu = \omega/2\pi$ and amplitude F_0 is applied to a dynamical system with *one* degree of freedom. The differential equation of the motion is then

$$m\ddot{\xi} + R\dot{\xi} + f\xi = F_0 \cos \omega t. \tag{9.4-1}$$

We shall find that there is considerable advantage in using complex quantities rather freely in this problem. Thus instead of using merely the *real* part of a complex force on the right side of (9.4-1), let us use $F_0e^{i\omega t}$, i.e., a complex force. The actual justification of this step may be made in the following manner. The oscillator equation may just as well be written

$$m\ddot{\xi}' + R\dot{\xi}' + f\xi' = F_0 \sin \omega t. \tag{9.4-2}$$

If now we multiply through $(9\cdot 4-2)$ by i and add to $(9\cdot 4-1)$ we obtain

$$m(\ddot{\xi} + i\ddot{\xi}') + R(\dot{\xi} + i\dot{\xi}') + f(\xi + i\xi') = F_0e^{i\omega t}$$
. (9.4-3)

This equation may be written

$$m\ddot{\xi} + R\dot{\xi} + f\xi = F_0 e^{i\omega t}, \qquad (9.4-4)$$

if we agree to let ξ now represent a complex displacement as in $(9\cdot4-3)$. On solving $(9\cdot4-4)$ we must ultimately take the real part of the resulting ξ (although, of course, the imaginary part would likewise describe equally well the behavior of the system under the influence of the force). The solution of $(9\cdot4-4)$ consists really of two parts, of which the first is that which has already been obtained in Sec. $9\cdot2$, while the second is that which on substitution will yield $F_0e^{i\omega t}$ on the right side instead of zero. The first is called the transient, for in any case it is more or less rapidly damped to zero. It is the second, or what may be called the steady state, solution in which we are interested here.

From the form of the right side, we are led at once to substitute

$$\xi = Ae^{i\omega t}, (9.4-5)$$

whence

$$\dot{\xi} = i\omega A e^{i\omega t}; \quad \ddot{\xi} = -\omega^2 A e^{i\omega t}. \tag{9.4-6}$$

On substitution there follows for A

$$A = \frac{F_0}{i\omega R + f - m\omega^2},\tag{9.4-7}$$

and the steady state solution is then

$$\xi = \frac{F_0^{i\omega t}}{i\omega R + f - m\omega^2}. (9.4-8)$$

The real part of this is now found at once by rationalizing the denominator. We multiply both numerator and denominator by $-i\omega R + f - m\omega^2$ and have

$$\xi = \frac{F_0[-i\omega R + f - m\omega^2]e^{i\omega t}}{(f - m\omega^2)^2 + \omega^2 R^2}.$$
 (9.4–9)

Therefore

$$\xi_{\text{real}} = \frac{F_0\{(f - m\omega^2)\cos\omega t + \omega R\sin\omega t\}}{\omega^2 R^2 + (f - m\omega^2)^2}.$$
 (9.4-10)

Let us introduce the angle α where

$$\cos\alpha = \frac{f - m\omega^2}{\sqrt{(f - m\omega^2)^2 + \omega^2 R^2}},$$

$$\sin \alpha = \frac{\omega R}{\sqrt{(f - m\omega^2)^2 + \omega^2 R^2}},$$

or

$$\tan \alpha = \frac{\omega R}{f - m\omega^2}.$$
 (9.4–11)

Our expression for the displacement then becomes

$$\xi_{\text{real}} = \frac{F_0 \cos (\omega t - \alpha)}{\sqrt{\omega^2 R^2 + (f - m\omega^2)^2}}.$$
 (9.4–12)

The angle α is the phase difference between the force and the dis-

placement. Before we discuss the displacement further it will be worth while to obtain the expression for ξ , the velocity, which is often more important in application than ξ itself. From $(9\cdot 4-6)$ we have

$$\dot{\xi} = i\omega\xi. \tag{9.4-13}$$

Hence from (9.4-8) there follows

$$\dot{\xi} = \frac{F_0 e^{i\omega t}}{R + i \left(m\omega - \frac{f}{\omega}\right)},\tag{9.4-14}$$

and if we evaluate the real part as above we get

$$\dot{\xi}_{\text{real}} = \frac{F_0 \cos (\omega t - \beta)}{\sqrt{R^2 + \left(m\omega - \frac{f}{\omega}\right)^2}},$$
 (9-4-15)

where

$$\tan \beta = \frac{m\omega - \frac{f}{\omega}}{R}, \qquad (9.4-16)$$

and β is the phase difference between the force and the velocity. Comparing (9.4-16) with (9.4-11) we see that

$$\tan \alpha = -\frac{1}{\tan \beta}.$$
 (9.4–17)

Hence

$$\alpha - \beta = \frac{\pi}{2} \cdot \tag{9.4-18}$$

This means that ξ and $\dot{\xi}$ differ in phase by $\pi/2$.

Now if we examine the expressions for ξ_{real} and ξ_{real} we note that the system is always forced to oscillate under the action of the periodic force, but that the amplitude of the resulting oscillation depends on the magnitudes of R, f, m, and ω , and of course may be very small indeed. On the other hand it may also be rather large. We note that if the characteristics of the system (R, f, m) are fixed, there is a value of the frequency of the applied force which makes $\sqrt{R^2 + (m\omega - f/\omega)^2}$ a minimum and hence the

amplitude of ξ_{real} a maximum. This happens for

$$\omega = \omega_0 = \sqrt{\frac{f}{m}}.$$
 (9.4–19)

When this occurs, the system is said to be in resonance with the force and the corresponding frequency

$$\nu_0 = \frac{\omega_0}{2\pi} = \frac{1}{2\pi} \sqrt{\frac{f}{m}} \tag{9.4-20}$$

is called the resonance frequency. For resonance then we have

$$\dot{\xi}_{\text{res.}} = \frac{F_0}{R} \cdot \cos \omega_0 t, \qquad (9.4-21)$$

since the phase difference β_0 for resonance is equal to zero. On the other hand the resonance displacement is

$$\xi_{\text{res.}} = \frac{F_0}{\omega_0 R} \cdot \cos\left(\omega_0 t - \frac{\pi}{2}\right).$$
 (9.4–22)

It is easily seen that the amplitude of the displacement is not quite a maximum for the resonance frequency. It reaches its maximum for that frequency for which $\sqrt{\omega^2 R^2 + (f - m\omega^2)^2}$ is a minimum. The usual test shows that this takes place for

$$\omega = \omega_1 = \sqrt{\frac{f}{m} - \frac{R^2}{2m^2}}.$$
 (9.4–23)

As a matter of fact if the damping is small the term $R^2/2m^2$ is negligible compared with f/m and $\omega_0 \doteq \omega_1$ approximately. Moreover, if we recall $(9\cdot 2-22)$ we remember that the so-called natural or free oscillation frequency is $1/2\pi \cdot \sqrt{f/m} - R^2/4m^2$ and this for small damping is approximately the same as the resonance frequency. We shall then attach most significance to the latter frequency.

Coming back to the phase we see that for resonance the force and the velocity of the system are exactly in phase, while the displacement and the force are out of phase by $\pi/2$. Any one who has pushed a swing will recall that the way to produce swings of large amplitude is to time the pushes so that they reach their maximum force when the displacement from equilibrium is smallest, not at the end of each swing.

The question of energy dissipation is an interesting one. As we have already shown in Sec. 9.3, the rate of dissipation is $R\xi^2$. Hence the *average* rate (in *time* of course) is

$$\overline{\dot{D}} = \frac{1}{\tau} \int_0^\tau R\dot{\xi}^2 dt$$

$$= \frac{F_0^2 R}{R^2 + \left(m\omega - \frac{f}{\omega}\right)^2} \cdot \frac{1}{\tau} \int_0^\tau \cos^2(\omega t - \beta) dt. \quad (9.4-24)$$

Now we have already shown (Sec. 9.3) that

$$\frac{1}{\tau} \int_0^\tau \cos^2 (\omega t - \beta) \, dt \doteq \frac{1}{2}, \tag{9.4-25}$$

if τ is much greater than a single period $P=\frac{2\pi}{\omega}$. Incidentally the student should show that on the other hand the average of $\sin (\omega t - \beta) \cos (\omega t - \beta)$ is zero. These results will be useful in the solution of special problems, and will be assumed from now on. We thus have

$$\overline{\vec{D}} = \frac{F_0^2 R}{2\left[R^2 + \left(m\omega - \frac{f}{\omega}\right)^2\right]}.$$
 (9.4–26)

This may be put into simpler form if we recall that

$$\cos \beta = \frac{R}{\sqrt{R^2 + \left(m\omega - \frac{f}{\omega}\right)^2}}.$$
 (9.4–27)

Therefore

$$\vec{D} = \frac{F_0^2}{2R} \cos^2 \beta.$$
 (9.4–28)

This form is particularly interesting because it is clear from it that when $\beta = 0$, i.e., when resonance ensues, \vec{D} is a maximum equal to

$$\overline{\dot{D}}_{\text{res.}} = \frac{F_0^2}{2R} \cdot \tag{9.4-28a}$$

On the other hand when $\beta = \pi/2$, i.e., when force and velocity are

out of phase (or force and displacement in phase or differing by π), $\vec{D} = 0$.

We have now to ask what connection $\overline{\dot{D}}$ has with the rate of contribution of energy by the force to the system. The latter will clearly be

$$\dot{W} = \xi F_0 \cos \omega t, \qquad (9.4-29)$$

that is, the product of the force by the velocity (which is the power). For the average we then have

$$\frac{\vec{W}}{\sqrt{R^2 + \left(m\omega - \frac{f}{\omega}\right)^2}} \frac{1}{\cos \omega t \cdot \cos (\omega t - \beta)}$$

$$= \frac{1}{2} \frac{F_0^2 \cos \beta}{\sqrt{R^2 + \left(m\omega - \frac{f}{\omega}\right)^2}}, \qquad (9.4-30)$$

whence from (9.4-27)

$$\overline{\dot{W}} = \frac{F_0^2}{2R} \cos^2 \beta = \overline{\dot{D}}. \tag{9.4-31}$$

In words, the rate at which the force contributes energy to the system is on the average just equal to the rate at which the system has its energy dissipated. This could indeed have been predicted from general considerations. The interesting thing is that when the system is in resonance with the force, the flow of energy *into* the system as well as the rate of dissipation by the system is a maximum.

The damping exercises a very important and characteristic effect on the resonance. Let us look into this. Writing (9.4–30) again we have

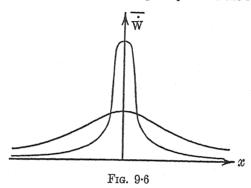
$$\overline{\dot{W}} = \frac{1}{2} \frac{F_0^2 R}{R^2 + \left(m\omega - \frac{f}{\omega}\right)^2}$$

$$= \frac{1}{2} \frac{F_0^2 R}{R^2 + m^2 \omega_0^2 x^2}, \qquad (9.4-32)$$

where we have put for convenience

$$\frac{\omega}{\omega_0} - \frac{\omega_0}{\omega} = x. \tag{9.4-33}$$

If now we plot $\overline{\dot{W}}$ as a function of x we get curves of the type indicated in the following figure (Fig. 9.6). When x = 0, $\omega = \omega_0$, and hence the value of $\overline{\dot{W}}$ at the origin represents the resonance or



maximum value. As ω deviates from ω_0 , $\overline{\dot{W}}$ falls off to either side. The rapidity of the dropping off determines the sharpness or broadness of the resonance. The latter is fixed by the value of $d\overline{\dot{W}}/dx$. Thus

$$\frac{d\overline{\dot{W}}}{dx} = -\frac{F_0^2 R m^2 \omega_0^2 x}{(R^2 + m^2 \omega_0^2 x^2)^2}.$$
 (9.4–34)

For x > 0, i.e., $\omega > \omega_0$, $\frac{d\overline{\dot{W}}}{dx}$ is negative and varies with R in such a way that the smaller R is, the greater is the absolute value of $\frac{d\overline{\dot{W}}}{dx}$.

On the other hand at $\omega = \omega_0$ for small R, $\overline{\dot{W}}$ is large. We may summarize by saying that for small damping factor the resonance peak is high but sharp, whereas for large damping factor the peak is low and blunt, i.e., the resonance is broad. This may be seen even more simply by a direct inspection of (9.4–32). The relation of the height of the peak to R can be seen by substituting x = 0, and the relative sharpness of the peaks can be inferred from the fact that when R is large one has to choose a larger value of x to

reduce the value of $\overline{\hat{W}}$ in the same proportion than when R is smaller.

Recent research in connection with mechanical oscillations has shown the value of introducing a notation borrowed from the theory of electrical oscillations. To go back to the expression for ξ [eq. $(9\cdot4-14)$] the amplitude of the velocity is controlled by the expression in the denominator, viz.,

$$R + i\left(m\omega - \frac{f}{\omega}\right). \tag{9.4-35}$$

This quantity will be denoted by Z and will be termed the *mechanical impedance* of the oscillating system. It is a complex quantity. If we split it into real and imaginary parts

$$Z = Z_1 + iZ_2, (9.4-36)$$

we have

$$Z_1 = R,$$
 (9.4–37)

and

$$Z_2 = m\omega - \frac{f}{\omega}. \tag{9.4-38}$$

The real part is called the *mechanical resistance*, while the imaginary part is called the *mechanical reactance*. The absolute value of the impedance is

$$|Z| = \sqrt{R^2 + \left(m\omega - \frac{f}{\omega}\right)^2}, \qquad (9.4-39)$$

so that we may write

$$\dot{\xi}_{\text{real}} = \frac{F_0}{|Z|} \cdot \cos (\omega t - \beta), \qquad (9.4-40)$$

i.e., for a given force the velocity amplitude varies inversely as the absolute value of the impedance. For the resonance case the impedance becomes real and reduces to the resistance. The significance of these terms will become clearer when we have discussed the electrical circuit analogy (cf. Sec. 9-6). But this notation has proved of value even in connection with mechanical problems.¹

¹ A very useful book on mechanical oscillations is that of J. P. Den Hartog, "Mechanical Vibrations" (McGraw-Hill, 2nd ed., New York, 1940).

9.5. The Acoustic Resonator as an Illustration of Oscillatory Motion. In order to emphasize the importance of the mechanics of simple oscillating systems, we may well pause here to consider some applications to physical problems. These are very numerous. We shall choose only a few typical illustrations.

Consider first the action of an acoustic resonator. This instrument (indicated diagrammatically in Fig. 9.7) consists usually of a hollow metal sphere with a small opening which may or may not have the form of a tubular neck. When a number of vibrating tuning forks of various frequencies are brought near the opening successively, it may be found that for one particular fork the reso-

nator appears to amplify the sound many times, while for the others little effect of this nature is observed. This phenomenon reminds one at once of the resonance condition discussed in Sec. 9-4 Hence we try to treat the action as a mechanical problem. The

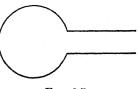


Fig. 9.7

student will find that this is a characteristic attitude taken by the physicist. Whenever an experimental phenomenon shows signs of similarity with some result of classical mechanics, the temptation is very strong to apply mechanical reasoning to the more exact description of the phenomenon. The justification of this process has come in the remarkable success which has so far attended it. It should be remarked, to be sure, that very recent physics has shown a tendency to get away from this program. However, we may well doubt that this tendency will ever be completely successful, so strong is the feeling of human minds for mechanical pictures. In any case we can hardly exaggerate in emphasizing the importance of mechanical methods in all the branches of physics.

Now if the acoustic resonator is a problem in mechanics we must decide what it is that *moves*. Here of course it is necessary to make an assumption, which however is not so very arbitrary. We postulate that the thing which moves is the *air* in the opening, and that there is a certain amount of it which moves *approximately* as a *whole*. If there is a neck (and this is the case we shall discuss here) we shall take as the *mass* of the moving air

$$m = \rho l S, \qquad (9.5-1)$$

where ρ is the density of the air, l the length of the neck and S its area of cross-section. If we denote the displacement of the air by ξ , the kinetic reaction then is

$$\rho lS\ddot{\xi}$$
. (9·5–2)

We must next look for the *stiffness* of the system. It seems reasonable that this will be found in the air in the chamber of the resonator which is alternately compressed and expanded as the air in the neck moves in and out, and hence acts like an elastic cushion for the latter. In order now to find the stiffness coefficient we must avail ourselves of a little elementary knowledge of elasticity and fluids (cf. Chapters X and XI). If a fluid of volume V changes its volume by an amount dV under the influence of pressure p (force per unit area — see Sec. 11·1 on hydrostatics) the necessary pressure is related to dV in the following way

$$\frac{p}{-\frac{dV}{V}} = k, \tag{9.5-3}$$

where k is the so-called volume elasticity or bulk modulus. The relation (9.5–3) is indeed the form Hooke's law (see Sec. 10.1) takes for fluids. Now in the problem under discussion, V will represent the volume of the resonator chamber and dV will be the change brought about in V by the motion of the "plug" of air in the neck. Hence we have

$$dV = -S\xi, (9.5-4)$$

for the volume change accompanying the air displacement ξ . The stiffness coefficient f has been defined previously to be the ratio of the force necessary to produce the displacement ξ to ξ itself. We therefore have

$$f=\frac{Sp}{\xi},$$

and using eqs. (9.5-3) and (9.5-4) this becomes

$$f = \frac{S^2k}{V}. (9.5-5)$$

We have finally to look for the damping factor. This is much the most difficult task. It might be thought that we shall find dissipation in the viscous resistance of the air itself and also in the friction at and absorption by the walls of the neck. These factors certainly enter but a more elaborate calculation than is appropriate for this book indicates that they alone are incompetent to account for the observed dissipation. Rather it develops that we must consider the radiation of sound energy from the opening of the resonator. From elementary physical principles it is clear that this radiation must exist, otherwise the amplification of the resonator at resonance could not be at all so evident. Its analytical calculation, however, is a rather difficult matter, since it is dependent on concepts connected with wave motion which we shall not discuss until Chapter X. However, it will do no harm to state here the result: the radiation produces a damping force of the following magnitude

$$R\dot{\xi} = \frac{\rho\omega^2 S^2}{2\pi c} \dot{\xi},\tag{9.5-6}$$

where c is the velocity of sound and $\omega = 2\pi\nu$, ν being the frequency of the radiated sound. The interesting point to note is that the damping force proves to be proportional to ξ as we assumed in the general mechanical case (Sec. 9.2). This furnishes a useful confirmation of our previous assumption. We may now write the equation of motion for the resonator as follows

$$\rho l S \xi + \frac{\rho \omega^2 S^2}{2\pi c} \dot{\xi} + \frac{S^2 k}{V} \xi = F_0 e^{i\omega t}, \qquad (9.5-7)$$

where $F_0e^{i\omega t}$ is the external force acting on the resonator due to the tuning fork, microphone or whatever other source of sound is used. We note at once the mathematical similarity between this equation and $(9\cdot 4-4)$. It is not necessary to carry through the solution. As an exercise the reader may show, for example, that the resonance frequency of the resonator is given approximately by

$$\frac{c}{2\pi}\sqrt{\frac{S}{lV}}. (9.5-8)$$

Here $c = \sqrt{k/\rho}$, the usual expression for the velocity of sound in a fluid (see Sec. 11.6). The form (9.5–8) would be strictly exact were it not for the damping which introduces a small correction term similar to that noted in the mechanical problem [eq. (9.2–22)].

Further details about the resonator, including for example the case of a resonator with *no* neck, and the general expression for the amplification are to be found in textbooks on acoustics. Our purpose here has been to emphasize the mechanical analogy, and it should be stressed that every theoretical development for the mechanical system in Sec. 9.4 has its counterpart in the acoustical system constituted by the resonator.

9.6. Electrical Oscillations. The concepts developed in Sec. 9.4 are of great value in electricity also, as we shall now proceed to illustrate with the simplest possible example. In developing it we shall assume merely the most elementary facts about electric currents such as the student should know from his first course in physics.

Let us suppose that we join a coil of heavy wire having inductance L but negligible resistance in series with a coil of wire of resistance R (wound back on itself to render its inductance

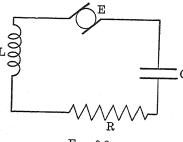


Fig. 9.8

negligible), and also insert in the circuit a condenser of capacity C. A source of alternating electromotive force $E = E_0 e^{i\omega t}$ is now placed across all three (see the diagrammatic sketch in Fig. 9-8). As in the case of the mechanical system we may consider this circuit a system having a certain "mass" (inertia), stiffness coefficient and

damping factor. The quantity which here corresponds to the mechanical displacement is the charge q. The kinetic reaction of the system is $L\ddot{q} = L\dot{I}$, where $\dot{q} = I$, the current flowing in the circuit; for $L\dot{I}$ is the back electromotive force due to the change of current (and hence of flux) through the inductance. We therefore look upon L as the effective mass of the circuit. An electromotive force E_c placed across the condenser of capacity C produces a charge $Q = E_c C$, and hence the stiffness coefficient, which is the ratio of force to corresponding displacement in the pure mechanical case, and which therefore would naturally be the

¹ See, for example, Stewart and Lindsay, "Acoustics" (D. Van Nostrand, New York, 1930), p. 47 ff.

ratio of electromotive force to electric displacement (or charge), in the present case is 1/C. Finally the damping force is $RI = R\dot{q}$, for this is of course the meaning of electrical resistance. The complete equation of motion for the forced oscillations of such a system is then

$$L\ddot{q} + R\dot{q} + \frac{q}{C} = E_0 e^{\imath \omega t}. \tag{9.6-1}$$

Once again the form is mathematically equivalent to eq. (9.4-4). The solution leads to the resonance frequency

$$\omega_0 = \sqrt{\frac{1}{LC}}, \qquad (9.6-2)$$

which is the frequency for which the current is a maximum. The frequency for which q is a maximum is, corresponding to (9.4-23),

$$\omega_1 = \sqrt{\frac{1}{LC} - \frac{R^2}{2L^2}},\tag{9.6-3}$$

which again does not usually differ very much from ω_0 , unless the resistance is very large. The absolute value of the *impedance* of the electrical circuit is

$$Z = \sqrt{R^2 + \left(L\omega - \frac{1}{\omega C}\right)^2}, \qquad (9.6-4)$$

and we note its analogy with the mechanical impedance (9.4–39). The reader is advised to find the electrical analogy for every important fact concerning the mechanical system, and refer for the physical significance to some text on electrical oscillations.¹

9.7. The Oscillator in Atomic Theory. Other important applications of harmonic oscillations are encountered in atomic theory. It will be recalled that in the atomic theory of the constitution of matter all bodies are assumed to be composed of small particles called *molecules* whose constituents in turn are atoms consisting of electrically charged particles called *electrons* and *nuclei*. We have

¹ See, for example, G. W. Pierce, "Electric Oscillations and Electric Waves" (McGraw-Hill, New York, 1920). See also L. Page and N. I. Adams, "Principles of Electricity" (D. Van Nostrand, New York, 1931). Consult also J. H. Morecroft, "Elements of Radio Communication" (Wiley, New York, 1929).

already had occasion in Sec. 3.9 to note the Bohr theory of the structure of hydrogen, the simplest of the atoms. Here the nucleus is a simple positively charged particle with most of the mass of the atom, and moving about it is a single electron negatively charged to the same magnitude as the positive charge of the nucleus, namely 4.8×10^{-10} electrostatic units. The nucleus of the hydrogen atom is called the *proton*. This is one of the constituents of the nuclei of other and more massive atoms. Besides protons, nuclei also contain *neutrons*, which are elementary particles possessing no electrical charge but having mass approximately equal to that of the proton. The positive charge of the nucleus is due to the protons alone, whereas the mass of the nucleus comes from the protons and neutrons together.

The atoms in a molecule are held together by mutual forces among their constituent particles. To assess these completely is indeed a complicated problem. However, to a first approximation much may be accomplished by lumping all the attractive forces between two atoms of a molecule into a single term varying with some appropriate power of the distance of separation, and doing likewise with the repulsive forces. This formulation seems reasonable from an examination of the individual electrostatic forces of attraction and repulsion between the individual particles of positive and negative sign. In general then we may write for the magnitude of the force between two atoms

$$F = a/r^p - b/r^q, (9.7-1)$$

where p and q are usually two different integers, and a and b are constants. Here r is the separation of the two atoms. Now from the discussion in Chapter VI (Sec. 6·1) it follows that we can reduce the motion of the two atoms subject only to their mutual force interaction to the motion of a single particle with the reduced mass and with displacement coördinate equal to r. Since the molecule must have some stability, it follows that for a certain value of r, say r_0 , F = 0. However in the neighborhood of this equilibrium position, certain motions are possible. Thus let $r = r_0 + x$, where x is very small compared with r_0 . Then

$$F = \frac{a}{(r_0 + x)^p} - \frac{b}{(r_0 + x)^q}$$

$$= \frac{a}{r_0^p} - \frac{b}{r_0^q} - \left(\frac{pa}{r_0^{p+1}} - \frac{qb}{r_0^{q+1}}\right)x + \cdots$$
 (9.7-2)

Now if x is small enough the expansion may be cut off at the second term. Moreover from the equilibrium condition just mentioned

$$a/r_0{}^p - b/r_0{}^q = 0. (9.7-3)$$

Hence the net force corresponding to displacement x from equilibrium is

$$F = -\left(\frac{pa}{r_0^{p+1}} - \frac{qb}{r_0^{q+1}}\right)x. \tag{9.7-4}$$

If the term in the parentheses is positive, it is clear that the motion of the two atoms will be simple harmonic along the line joining them and indeed with frequency

$$\nu = \frac{1}{2\pi} \sqrt{\frac{(pa/r_0^{p+1} - qb/r_0^{q+1})}{m}},$$
 (9.7-5)

where m is the reduced mass. If the term in the parentheses were to be negative, we should not get harmonic motion, but in this case it is clear that the equilibrium at $r = r_0$ could not be stable.

The whole analysis of this chapter may be applied to the oscillations of charged particles like those mentioned in the preceding paragraph. Many actual cases have been studied ideally by supposing that the electrons attached to atoms vibrate as linear In particular they may be considered to be subjected to damping forces which must be compensated for by external periodic forces to keep the vibrations going. The latter will in this case clearly be electric fields which, for example, may be due to radiation falling on the substance which is composed of the charged particles. If the frequency of an imposed field is equal to the natural frequency of the oscillator we shall expect all the phenomena of resonance, with resulting absorption of the incident This is properly the subject of investigation by the electromagnetic theory of optics, which is too imposing a region for us to enter here. The student, however, will find there ample illustration of the properties of oscillations discussed from the mechanical point of view in this chapter.

The following interesting question may be raised: since the Bohr theory is based on the idea that not all the possible mechanical motions of the electrons in an atom are allowed but only certain ones, specifically picked out by a process of quantization (Sec. 3-9), should not the same treatment be applied to the harmonic oscillator? The answer to this question is in the affirmative.

The total energy of such an oscillator assumed for simplicity to be undamped [Sec. 9·3, eqs. (9·3–6) and (9·3–7)] is given by

$$U = T + V = 2m\pi^2\nu^2 A^2, (9.7-6)$$

where m is the mass, ν the frequency and A the amplitude of the oscillator. As far as classical mechanics is concerned, for a given oscillator this may have any value whatever, depending on the choice of amplitude. Of course in most mechanical problems the amplitude must not be too great if the oscillation is to be simple harmonic at all — but within this range all values are mechanically possible. However, the quantum theory prescribes that the oscillator may exist only in those states of motion for which the space integral of the momentum over a complete cycle is equal to an integral multiple of the fundamental constant of action, viz., h, the so-called Planck's constant. We may write the quantum condition (cf. 3-9-1) in the case of linear oscillatory motion thus:

$$2\int_{-A}^{+A} m\dot{x} \, dx = nh, \tag{9.7-7}$$

when n is an integer. The factor 2 enters since the integration must be extended from x = +A to x = -A and back to x = +A. Since

$$x = A \sin (2\pi \nu t + \epsilon)$$

we have

$$\dot{x} = 2\pi\nu A \cos(2\pi\nu t + \epsilon)$$

$$= 2\pi\nu A \sqrt{1 - \frac{x^2}{A^2}}, \qquad (9.7-8)$$

and therefore on substituting into the integral the result is

$$4\pi\nu m \int_{-A}^{+A} \sqrt{A^2 - x^2} \, dx = 2\pi^2 \nu m A^2 = nh. \qquad (9.7-9)$$

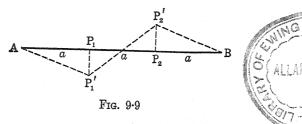
This at once yields for the allowed energy values

$$U_n = 2\pi^2 m \nu^2 A^2 = nh\nu. (9.7-10)$$

It then develops that the permitted energy values for the simple harmonic oscillator are integral multiples of the quantity $h\nu$, which thus appears as a fundamental unit of energy in atomic theory. It is called the *energy quantum* and plays a role of over-

whelming importance in the modern theory of atomic structure. What we wish to emphasize at this place, however, is the fact that the imposition of the quantum condition leads to the introduction of a certain discreteness into the problem of oscillator motion, a discreteness which contrasts markedly with the essential continuity of the possible states of motion in classical mechanics. The necessity structure must account for the essentially discontinuous nature of atomic phenomena. It is pertinent to remark further that the method of quantization used in the Bohr theory has been replaced a brief survey of this is presented in Sec. 12.4, and the problem of the oscillator is there worked out on the basis of the new view.

9.8. Oscillations of a System with Several Degrees of Freedom-So far in our study of oscillations we have restricted ourselves to systems with one degree of freedom. There are, however, many interesting features associated with systems of two or more degrees of freedom. The following is a simple example. Consider



the case of a horizontal string AB (Fig. 9.9) of negligible mass loaded with two mass particles at P_1 and P_2 . We let $\overline{AP_1} = \overline{P_2B} = \overline{P_1P_2} = a$, for simplicity. The mass of the particle at P_1 is m_1 while that at P_2 is m_2 . The ends A and B are fastened, the string being stretched with tension τ .² The particles are displaced slightly from their equilibrium position and then let go; we have to determine their subsequent motion, assumed to lie in the vertical plane through the string. Let the displacements at any instant

¹ The student may here refer again to Ruark and Urey, "Atoms, Molecules, and Quanta," for further information on the quantum theory.

² Actually, of course, the weight of the particles will keep the string from maintaining an exact horizontal line in the equilibrium position. Nevertheless if the tension is sufficiently great the deviation is for our purposes negligible.

be ξ_1 and ξ_2 respectively. We have then to write down the equations of motion. Consulting the figure, we see that the restoring force on the first particle, assuming the tension is unaltered by the displacement, is

$$\tau \left[\frac{\xi_1}{a} - \frac{\xi_2 - \xi_1}{a} \right] = \frac{\tau}{a} [2\xi_1 - \xi_2], \tag{9.8-1}$$

while that on the second is

$$\tau \left[\frac{\xi_2}{a} + \frac{\xi_2 - \xi_1}{a} \right] = \frac{\tau}{a} [2\xi_2 - \xi_1]. \tag{9.8-2}$$

The equations of motion when the system is free are therefore¹

$$m_{1}\ddot{\xi}_{1} + \frac{\tau}{a} (2\xi_{1} - \xi_{2}) = 0,$$

$$m_{2}\ddot{\xi}_{2} + \frac{\tau}{a} (2\xi_{2} - \xi_{1}) = 0.$$
(9.8-3)

To solve these equations, let us introduce the "operators"

and

$$D_{1} \cdot = \left(m_{1} \frac{d^{2}}{dt^{2}} + \frac{2\tau}{a}\right) \cdot ,$$

$$D_{2} \cdot = \left(m_{2} \frac{d^{2}}{dt^{2}} + \frac{2\tau}{a}\right) \cdot ,$$

$$(9.8-4)$$

the meaning of these being that when D_1 operates on ξ_1 we get

and similarly $D_1\xi_1 = m_1\ddot{\xi}_1 + \frac{2\tau}{a}\xi_1,$ $D_2\xi_2 = m_2\ddot{\xi}_2 + \frac{2\tau}{a}\xi_2.$ (9.8-5)

If now we write the equations in this form they become

$$D_{1}\xi_{1} - \frac{\tau}{a}\xi_{2} = 0,$$

$$D_{2}\xi_{2} - \frac{\tau}{a}\xi_{1} = 0.$$
(9.8-6)

¹ It is to be noted that we are here neglecting dissipation. This is solely for the sake of making the analysis simple. For more details consult Rayleigh, "Theory of Sound" (Macmillan, New York, 1929), Vol. I, para. 120 ff.

Let us multiply through the first equation by τ/a and operate on the second by D_1 . We then have

$$\frac{\tau}{a}D_1\xi_1-\frac{\tau^2}{a^2}\xi_2=0,$$

$$D_1 D_2 \xi_2 - \frac{\tau}{a} D_1 \xi_1 = 0.$$

Adding these two equations yields

$$D_1 D_2 \xi_2 - \frac{\tau^2}{a^2} \xi_2 = 0, (9.8-7)$$

so that ξ_1 has been eliminated. The resulting equation for ξ_2 is then

$$m_1 m_2 \ddot{\xi}_2 + \frac{2\tau}{a} (m_1 + m_2) \ddot{\xi}_2 + \frac{3\tau^2}{a^2} \xi_2 = 0.$$
 (9.8-8)

There is a precisely similar equation for ξ_1 , viz.,

$$m_1 m_2 \ddot{\xi}_1 + \frac{2\tau}{a} (m_1 + m_2) \ddot{\xi}_1 + \frac{3\tau^2}{a^2} \xi_1 = 0.$$
 (9.8-9)

Let us try the solution

$$\xi_1 = Ae^{\theta t}. \tag{9.8-10}$$

Substitution into eq. (9.8-9) yields

$$m_1 m_2 \theta^4 + \frac{2\tau}{a} (m_1 + m_2)\theta^2 + \frac{3\tau^2}{a^2} = 0.$$
 (9.8–11)

Solution for θ^2 gives

$$\theta^{2} = \frac{-\frac{2\tau}{a} (m_{1} + m_{2}) \pm \sqrt{\frac{4\tau^{2}}{a^{2}} (m_{1} + m_{2})^{2} - \frac{12\tau^{2}}{a^{2}} m_{1} m_{2}}}{2m_{1}m_{2}} \cdot (9.8-12)$$

There are thus four values of θ , as there should be since the equation (9.8-9) is of the fourth order. They are rather hard to handle in the general case where $m_1 \neq m_2$. To simplify the subsequent

solution, let us at this point assume $m_1 = m_2 = m$. We then have from (9.8-12)

$$\theta^2 = \frac{-\frac{4\tau m}{a} \pm \frac{2\tau m}{a}}{2m^2} = \frac{-3\tau}{ma} \text{ or } \frac{-\tau}{ma}.$$
 (9.8–13)

Hence

$$\theta = \pm i \sqrt{\frac{3\tau}{ma}} \quad \text{or} \quad \pm i \sqrt{\frac{\tau}{ma}},$$
 (9.8–14)

and the solutions for both ξ_1 and ξ_2 take the form

$$\xi_1 = A_1 e^{i\sqrt{3\tau/ma} \cdot t} + A_1' e^{-i\sqrt{3\tau/ma} \cdot t}$$

$$+ B_1 e^{i\sqrt{\tau/ma} \cdot t} + B_1' e^{-i\sqrt{\tau/ma} \cdot t}, \qquad (9.8-15)$$

$$\xi_2 = A_2 e^{i\sqrt{3\tau/ma} \cdot t} + A_2' e^{-i\sqrt{3\tau/ma} \cdot t}$$

$$+ B_2 e^{i\sqrt{\tau/ma} \cdot t} + B_2' e^{-i\sqrt{\tau/ma} \cdot t}.$$

$$(9.8-16)$$

If we reduce to real form (as in Sec. 9.1) we have finally

$$\xi_{1} = a_{1} \cos \left(\sqrt{\frac{3\tau}{ma}} t + \epsilon_{1} \right) + b_{1} \cos \left(\sqrt{\frac{\tau}{ma}} t + \epsilon_{1}' \right),$$

$$\xi_{2} = a_{2} \cos \left(\sqrt{\frac{3\tau}{ma}} t + \epsilon_{2} \right) + b_{2} \cos \left(\sqrt{\frac{\tau}{ma}} t + \epsilon_{2}' \right),$$

$$(9.8-17)$$

where the amplitudes a_1 , b_1 , a_2 , b_2 , and the phases ϵ_1 , ϵ_2 , ϵ_1' , ϵ_2' are not all independent. In fact, since we started originally with two second order equations, we must have only *four* independent arbitrary constants. If we substitute (9.8-17) back into (9.8-3) with $m_1 = m_2 = m$, we obtain after some algebraic manipulation the following simple results:

$$a_2 = a_1, \quad b_2 = -b_1, \quad \epsilon_2 = \epsilon_1, \quad \epsilon_2' = \epsilon_1'.$$
 (9.8–18)

Physically we may interpret (9.8–17) as follows. Both particles have a motion which is the result of compounding two simple harmonic oscillations with frequencies

$$\nu_1 = \frac{1}{2\pi} \sqrt{\frac{\tau}{ma}}, \quad \nu_2 = \frac{1}{2\pi} \sqrt{\frac{3\tau}{ma}}.$$
 (9.8–19)

These two frequencies may be called the characteristic frequencies

of the system, and the latter may be said to have two distinct modes of oscillation. It may be remarked that since the two frequencies ν_1 and ν_2 are incommensurable the resulting motion will not in general be *periodic*, i.e., unless one of the two amplitudes happens to be zero. The four phase constants and amplitudes must be determined by the initial conditions of the motion. We have already said enough about these conditions in our discussion of the case of one degree of freedom to render it unnecessary to enter on details here. We should, however, stress the fact that the final solution for both particles is obtained by adding or superposing separate terms which are themselves solutions. reason for the possibility of this is to be found in the linearity of the original equations of motion (9.8-8, 9), i.e., in the fact that they contain no second or higher powers of the ξ 's or any of their derivatives. It is the special characteristic of a linear equation that the sum of two or more solutions is a solution (see Sec. 10.4, where this is demonstrated). The differential equations of oscillatory motion for small disturbances from equilibrium are always linear equations, as may well be suspected from the simple illustration treated here. Hence the solution in each case will consist of a superposition of special modes of oscillation. known as the superposition principle. We shall have other illustrations of it later.

As might be supposed the only modification made in this result by the introduction of damping factors will be to make the amplitude of each mode contain a term of the form $e^{-\lambda t}$ (λ being real) which damps out the resulting oscillations as time progresses.

The general problem (of which ours is a special case) of the motion of n equal mass particles attached at equal intervals along a finite string stretched with tension τ is of great historical interest for it was first solved by Lagrange in his famous treatise "Mechanique Analytique," using the powerful general methods introduced by him into mechanics. We merely note that as one might expect from our special case the motion of each particle is the superposition of n simple harmonic motions each with a characteristic frequency. There are then n distinct modes of oscillation.

As other interesting physical illustrations of the oscillations of a system of two or more degrees of freedom, we may note: (1) the vibrations of coupled resonators, i.e., two resonators joined in series; (2) coupled electrical circuits with resistance, induc-

tance and capacity (here the damping can rarely be neglected), such as are encountered in every radio set; and (3) coupled oscillations of charged electric particles, such as the ions in a crystal. In each case the method of attack outlined above may be employed and the results interpreted analogously to the mechanical case.

In the case of electrical oscillations an operational method of attack originated by Oliver Heaviside has been of great utility • and can of course be applied to oscillating systems in general.

PROBLEMS

- 1. Compare the *space* average of the kinetic energy of a simple harmonic oscillator over the space interval from x = 0 (the equilibrium position) to x = A, where A is the amplitude, with the corresponding average of the potential energy and comment on the result.
- 2. In the case of a simple harmonic oscillator derive an expression for the fractional part of a whole period which it spends in a small interval Δx at distance x from its equilibrium position. Construct curves plotting the frac-

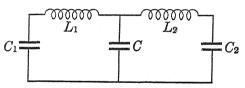
tional time as a function of x for several different amplitudes and comment on their physical significance.

- 3. The schematic diagram shows a mass m able to move vertically under the action of two massless springs with stiffness coefficients f_1 and f_2 respectively. Find the resonance frequency of the system if it performs simple harmonic oscillations. Find the percentage change in the resonance frequency if f_1 and f_2 change by Δf_1 and Δf_2 respectively (where $|\Delta f_1| \ll f_1$ and $|\Delta f_2| \ll f_2$).
- 4. A mass of 50 grams is attached to the center of a horizontal string 20 cm long stretched with a tension equivalent to 2000 gram wt. Assuming that the mass of the string is negligible, find the resonance frequency of the equivalent stiffness of the system. If it is observed that the amplitude of vibration is reduced in the ratio 1 to e (Naperian base) in 5 seconds, find the damping coefficient and the logarithmic decrement.
- 5. Referring to Problem 4, use the damping coefficient to calculate the natural frequency of oscillation and the frequency for maximum amplitude.
- ¹ For details see, for example: R. V. Churchill, "Modern Operational Mathematics in Engineering" (McGraw-Hill, New York, 1944); H. S. Carslaw and J. C. Jaeger, "Operational Methods in Applied Mathematics" (Oxford Univ. Press, New York, 1941); N. W. McLachlan, "Modern Operational Calculus" (Macmillan, London, 1948).

Calculate the mechanical impedance of the system at (a) the resonance frequency and (b) a forced frequency of 50 cycles/sec.

- 6. A certain tuning fork of frequency 256 cycles has its amplitude diminished in the ratio 1/e in about 5900 cycles. Calculate the decay modulus and the logarithmic decrement. Find the value of the interval ω/ω_0 [see eq. (9.4–33)] in which the intensity of resonance (i.e., \overline{W}) falls to one-half its maximum value. If the tuning fork is placed before the opening of an air resonator of equal natural frequency, the number of cycles to bring about the amplitude reduction specified is 3300. Discuss the physical reason for this and calculate ω/ω_0 for this case also. Plot representative curves for \overline{W} in the two cases.
- 7. A Helmholtz resonator of natural frequency 256 cycles/sec is in the form of a sphere with volume 1053 cm³. If the diameter of the opening is 3.02 cm, calculate the effective length of the neck on the opening. of the word "effective" implies that there may actually be no neck present, yet there is a plug of air vibrating in the opening and the length in question may be considered to give the dimension of this plug. See Stewart and
- 8. If the resonator in Problem 7 is excited by a tuning fork of frequency 256 which exerts a maximum force of 25 dynes at the opening, compute the maximum displacement velocity of the air in the opening. Do the same for the case of a fork of frequency 512 exerting the same force.
- 9. A mechanical vibrating system has a mass of 100 grams and a stiffness factor of 1.6×10^7 dynes/cm and a damping coefficient of 200 dynes sec/cm. Calculate the resonance and natural frequencies and find the percentage difference.
- 10. Find the decay modulus and logarithmic decrement for the system specified in Problem 9. If the system is excited by a force of frequency 100 cycles/sec with a maximum of 10 grams wt, calculate the average rate of
- 11. Find the expression for the natural frequency of the oscillations of an electric circuit made up of resistance R, inductance L, and capacitance C in
- 12. The cosine of the phase angle between impressed electromotive force and current in an electric circuit containing resistance, inductance, and capacitance in series is called the power factor. Suggest a reason for this name and calculate the power factor in a case where the resistance is 1 ohm, the inductance is 4 millihenrys (a henry is the inductance of a circuit in which an impressed e.m.f. of one volt causes the current to grow at the rate of one ampere per sec) and the capacitance is 3 microfarads (a farad is the capacitance of a condenser for which a charge of one coulomb corresponds to a difference of potential of one volt). If a periodic e.m.f. of 100 volts with frequency 60 cycles/sec is impressed on this circuit, calculate the average rate of dissipation of energy by

13. The figure shows two electric circuits coupled together. Find the characteristic resonance frequencies of the system.

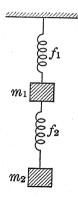


14. Consider two masses m_1 and m_2 joined together by a light rod whose longitudinal elasticity is so very large compared with its mass that the latter may be neglected. If the masses are displaced in the direction of the rod, set up and solve the equations of motion of the system. Show that the frequency of the motion is

$$\nu = \frac{1}{2\pi} \sqrt{\frac{f(m_1 + m_2)}{m_1 m_2}},$$

where f is the stiffness coefficient of the rod (i.e., f = YS/l where Y is Young's modulus, S = cross-sectional area of the rod, and l is its length). What will

be the effect on the frequency if the mass of the rod (say m_r) is not neglected? Why in this problem is there but one natural frequency instead of two?



15. In the schematic diagram m_1 and m_2 are masses suspended by springs with stiffness f_1 and f_2 respectively. Assume that a harmonic force of frequency $\omega/2\pi$ acts on m_1 . Find the steady-state displacement of each mass. What happens when the resonance frequency of m_2 by itself is equal to $\omega/2\pi$? (This is the principle of the Frahm vibration absorber.)

16. According to classical electrodynamics an accelerated electron radiates energy at the rate $\frac{2}{3}e^2f^2/c^3$ where e is the charge on the electron, e the velocity of light, and f is the acceleration. Assuming an electron in simple har-

monic motion with frequency ν calculate the amount of energy radiated during a single period. Find the number of periods it would take to reduce the energy of the motion to one-half its original value.

17. The string of a simple pendulum is gradually shortened at a rate very slow compared with the period of the pendulum. Calculate the average work that is done in the shortening over a single period. Derive the expressions for the kinetic and potential energies of the pendulum and show that during the shortening process the value of E/ν remains unaltered, E being the total energy. Prove that $E/\nu = \int p \, dx$, where p is the momentum of the pendulum bob. (E/ν) is called an adiabatic invariant, and its invariance in such a process as indicated is of great importance in atomic theory. N.B. $\int p \, dx$ is the integral taken over one complete cycle of the motion.)

18. In the diatomic molecule of hydrogen chloride the masses of the hydrogen and chlorine atoms are respectively:

 $m_{\rm H} = 1.66 \times 10^{-24} \, {\rm gram}$ and $m_{\rm Cl} = 5.9 \times 10^{-23} \, {\rm gram}$.

The vibration frequency has been found to be 8.721×10^{13} cycles/sec. What is the effective stiffness of the system? If the force curve is fitted with a function like that shown in (9.7-1) with p=2 and q=3, and it is known that $r_0=1.28\times 10^8$ cm, find the constants a and b. Find the force associated with a displacement 10^{-9} cm from equilibrium.

CHAPTER X

DEFORMABLE BODIES AND WAVE MOTION

10.1. Strain and Stress. Hooke's Law. In our work so far we have studied the motion of particles, both singly and in groups. But our only treatment of the behavior of large scale solids such as are actually met in physical experiments has been the discussion of the motion of rigid bodies, the constitution of which has been supposed to be such that the distance between any two points of the body remains forever unaltered no matter what forces act on it. This means that such a body can never be deformed. It was pointed out in that connection that the rigid body is a highly idealized concept, even though valuable for the description of many phenomena. However, there are enough illustrations of the deformation of bodies, changes in size and shape, to make it necessary to deal with these problems at some length and by methods differing in some important respects from those already used in treating particle motion. For we shall now effectively consider a large scale body as a continuous medium, rather than as made up of discrete particles, and the continuity thus assumed will play an important role in our theoretical considerations. have occasion to note the great theoretical importance of the distinction between particle and medium motion.

For the moment, however, let us confine our attention to the possible changes in size and shape which a finite portion of a continuous material medium can undergo. Such alterations are called *strains*, and we may classify them under a relatively few simple types.

Perhaps the simplest is change in volume without change of shape. Suppose we have a body of volume V and by the application of appropriate forces the volume is changed by an amount ΔV . This may be either positive or negative, corresponding to an expansive or compressive strain. It is advisable to fix upon a definite measure of volume strain, and for convenience we shall agree to call

$$\delta_V = \frac{\Delta V}{V} \tag{10.1-1}$$

the measure of the volume strain. It is the change in volume per unit volume, and implies the possibility of a change in the average density, ρ , of the body. It need not be the same for all parts of the body; if we divide the body into volume elements the δ_V for each element will not necessarily be the same. In this case the strain is said to be non-uniform. For the present we shall concern ourselves primarily with uniform strains. Since the mass m of the body is not changed, and since

$$m = \rho V, \tag{10.1-2}$$

we can also write for the volume strain

$$\delta_V = \frac{\Delta V}{V} = -\frac{\Delta \rho}{\rho}, \qquad (10 \cdot 1 - 3)$$

a relation which will later be useful. There is a special case of volume strain which we ought to notice, viz., a *linear strain*. Suppose the distance between any two points in the unstrained body is l. Under strain the distance becomes $l + \Delta l$. The linear strain is then defined by

$$\delta_l = \frac{\Delta l}{l} \cdot \tag{10.1-4}$$

It is seen at once that there is a relation between linear and volume strains. For consider a cube of side l. The volume is $V = l^3$. If the alterations are small enough so that they may be treated approximately as differentials we have

$$\Delta V = 3l^2 \, \Delta l, \qquad (10 \cdot 1 - 5)$$

so that on dividing through by V there results

$$\delta_V = 3\delta_l. \tag{10.1-6}$$

Hence a small uniform volume strain δ_V may be considered as replaced by three equal linear strains in three mutually perpendicular directions of magnitude $\delta_V/3$. It must be emphasized again that this implies strains small enough so that δ_V^2 and δ_V^3 etc., may be neglected compared with δ_V . It is with strains of this order of magnitude that we shall deal exclusively in this chapter. Liquids and gases can undergo volume strain as well as solids. On the other hand linear strain is confined to solids.

300 DEFORMABLE BODIES AND WAVE MOTION

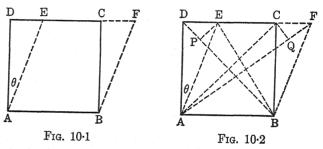
A particularly interesting case of linear strain is that of a long cylindrical wire. If the wire is stretched in the direction of its length a positive linear strain ensues. There is a concomitant lateral strain, i.e., the diameter of the wire is decreased. If we denote the lateral strain by $\delta_D = \Delta D/D$, where D is the diameter of the wire, then

$$\sigma = \frac{\delta_D}{\delta_l} \tag{10.1-7}$$

is known as Poisson's ratio. Its significance will be discussed in the next section.

Volume and linear strains are often grouped together under the caption, dilatational strains.

The last type of strain to be discussed is the so-called *shear*. Consider the cross section of a cube, *ABCD* (Fig. 10·1). Holding



fixed the plane of which \overline{AB} is the trace, allow all the other parallel planes in the cube to be shifted toward the right parallel to themselves, each one an amount proportional to its distance from AB, so that the square cross section ABCD becomes the rhombus ABFE. In this case the cube is said to have been sheared and the strain is known as a shearing strain. More particularly it is a uniform shearing strain. Now what shall be chosen as the measure of the strain? It is customary to take for this the value of the angle θ , which for small shears will be approximately equal to $\overline{DE}/\overline{AD}$, i.e., $\tan\theta$. An important thing about a shear is that no volume change is involved, it being wholly a change in shape, as far as the body as a whole is concerned. Of course length alterations are involved, and it is clear from an inspection of Fig. 10·1 that all lengths originally parallel to the diagonal \overline{AC} are increased while all those originally parallel to the diagonal \overline{BD} are decreased

by the shear. As simple experience amply indicates it is not possible to subject most liquids and gases to shearing strain. Nevertheless the shear does have meaning for very viscous liquids as well as for solids (cf. Sec. 11·1).

We shall now prove a significant theorem about shearing strains. For this purpose consult Fig. 10·2, which is like Fig. 10·1, except that the diagonals \overline{AC} and \overline{BD} have been drawn, as well as the new diagonals \overline{AF} and \overline{BE} . Let us draw \overline{EP} perpendicular to \overline{BD} , and \overline{CQ} perpendicular to \overline{AF} . Now since the shear is small the angle CFQ is still approximately 45°, and hence, approximately,

$$\overline{QF} = \frac{\overline{CF}}{\sqrt{2}}, \qquad (10.1-8)$$

while to the same approximation

$$\overline{AC} = \overline{AQ}.$$

Hence the positive linear strain along the one diagonal is

$$\frac{\overline{QF}}{\overline{AC}} = \frac{\frac{\overline{CF}}{\sqrt{2}}}{\sqrt{2}\overline{AD}} = \frac{1}{2}\frac{\overline{DE}}{\overline{AD}} = \frac{\theta}{2}, \qquad (10.1-9)$$

since θ is so small that, as we have said before, $\theta \doteq \tan \theta$. We can show similarly that the negative linear strain along the other diagonal has the same magnitude, viz., approximately,

$$\frac{\overline{DP}}{\overline{BD}} = \frac{\theta}{2}.$$
 (10·1–10)

We have thus shown that a shear strain θ may be considered as equivalent to a positive linear strain $\theta/2$ at right angles to a negative linear strain $\theta/2$, both being at 45° to the sheared planes.

We have been considering so far only the geometry of deformation. The strains just defined are always associated with forces which are applied to the body in question, and the equal and opposite reaction forces with which the body acts on the external influence. These forces may always be considered as distributed over some area or areas of the body, and the force per unit area is denominated the stress. Corresponding then to different types of strain, we have different types of stress. For example, with volume strain will be associated what may be called compressive stress

(or sometimes hydrostatic stress or pressure. See Sec. 11·1). Suppose we consider once more a cube of side l. If on every unit of area of every face of the cube the same force acts normal to the surface, it is said to be under the action of a compressive or expansive stress, depending on the direction of the force with respect to

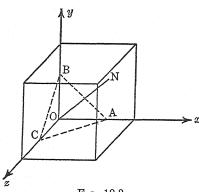


Fig. 10.3

the outward drawn normal. As we shall have occasion to note in Sec. 11·1, the former is true when a body is submerged in a fluid — in so far as the action of the fluid on the body is concerned.

The idea of compressive stress gains increased significance from the fact that if we draw any plane whatever through the cube just considered, the normal stress on this plane will be the same as that

on the faces of the cube. To show this consider (Fig. 10·3) a cube placed with one vertex at O, the origin of a system of rectangular coördinates. Let the plane be the one which has intercepts \overline{OA} , \overline{OB} , \overline{OC} on the x, y, z axes respectively. We shall represent its equation by

$$ax + by + cz = d,$$
 (10·1-11)

where

$$\overline{OA} = \frac{d}{a}, \quad \overline{OB} = \frac{d}{b}, \quad \overline{OC} = \frac{d}{c}.$$
 (10·1–12)

 \overline{ON} represents the normal to the plane through the origin. We know from analytic geometry that the direction cosines of the normal are

$$\frac{a}{\sqrt{a^2+b^2+c^2}}$$
, $\frac{b}{\sqrt{a^2+b^2+c^2}}$, $\frac{c}{\sqrt{a^2+b^2+c^2}}$. (10·1–13)

We wish to calculate the stress on the plane ABC. Let the stress on each face of the cube be denoted by p. Then the force on the area OBC in the yz plane is in magnitude pS_{OBC} (where S_{OBC} stands for the magnitude of this area), and its direction is along the positive x axis. This force also acts on the area ABC. Its com-

ponent normal to the plane will then be (recalling the significance of the direction cosine)

$$pS_{OBC} \frac{a}{\sqrt{a^2 + b^2 + c^2}}$$
 (10·1–14)

If now we proceed with the other faces similarly we find for the total normal *force* on the plane ABC due to the compressive stress on all the faces of the cube,

$$\frac{p}{\sqrt{a^2 + b^2 + c^2}} [aS_{OBC} + bS_{OAC} + cS_{OAB}]. \quad (10.1-15)$$

But to find the normal stress we must divide the force by the area. Hence the stress is

$$\frac{p}{\sqrt{a^2 + b^2 + c^2}} \left[\frac{aS_{OBC} + bS_{OAC} + cS_{OAB}}{S_{ABC}} \right] \cdot (10 \cdot 1 - 16)$$

Now, however, since S_{OBC} is the projection of S_{ABC} on the yz plane

$$\frac{S_{OBC}}{S_{ABC}} = \frac{a}{\sqrt{a^2 + b^2 + c^2}},$$
 (10·1–17)

and similarly for the others. Hence the normal stress reduces to

$$\frac{p}{\sqrt{a^2 + b^2 + c^2}} \left[\frac{a^2 + b^2 + c^2}{\sqrt{a^2 + b^2 + c^2}} \right] = p, \quad (10.1-18)$$

as was to be proved.

Corresponding to linear strain we may set tensile stress, which is the force per unit area acting on every cross section perpendicular to the direction of the strain. Finally corresponding to shearing strain we have shearing stress, which is the tangential force per unit area parallel to the shearing planes.

The next and very important question is: What is the relation between each strain and the corresponding stress? The answer to this is found in the celebrated law of Hooke. For a great many bodies actually observed in nature the ratio of the stress to the strain is approximately a constant for a certain range of variation of the stress. Such bodies are termed elastic bodies and the study of their behavior is often called elasticity, for when the external stress is removed these bodies resume more or less approximately their original size and shape in distinction to plastic bodies, which to a large extent retain their deformation. It must be emphasized

304 DEFORMABLE BODIES AND WAVE MOTION

that there is no hard and fast line between elasticity and plasticity, for a body may be elastic with respect to a certain type of stress and plastic with respect to another. For example, a viscous fluid is plastic with respect to shearing stresses but approximately elastic with respect to compressive stresses. We shall have occasion in Sec. 10·3 to discuss this aspect of the nature of deformable bodies in a little more detail. (Recall also the discussion of relaxation phenomena in Sec. 9·2.)

We now wish to introduce the expressions for Hooke's law for the different types of stress-strain combinations. First consider volume strains. Denoting as before the compressive stress by p, with δ_V as the corresponding volume strain, it is found that the ratio p/δ_V has a value which for many bodies over a wide range of values of p is constant. Thus we set

$$\frac{p}{\delta_{v}} = -k. \tag{10.1-19}$$

TABLE OF VALUES OF ELASTIC MODULI¹

Substance	Y Young's Modulus (dynes/cm²)	$\mu \\ \mathrm{Shear} \\ \mathrm{Modulus} \\ \mathrm{(dynes/cm^2)}$	k Bulk Modulus (dynes/cm²)	σ Poisson's Ratio
Aluminum. Brass (cold rolled). Copper. Glass (window). Glycerine. Gold. Granite. Lucite. Mercury Nickel. Polystyrene. Silver Silver Steel. Water. Wood (mean).	$\begin{array}{c} 7 \times 10^{11} \\ 9 \\ 12 \\ 7 \\ \hline 8 \\ 5 \\ 5.6 \\ \hline 20 \\ 3.9 \\ 7.5 \\ \hline 18 \\ \hline 0.5 \end{array}$	2.5 × 10 ¹¹ 3.5 4.3 2.7 3.0 1.8 2.15 7.1 1.5 2.7 7	$\begin{array}{c} 12 \times 10^{11} \\ 7 \\ 20 \\ 5.8 \\ 0.50 \\ 8.3 \\ 5 \\ 4.9 \\ 0.27 \\ 37 \\ 3.6 \\ 11.5 \\ 25 \\ 0.21 \\ - \end{array}$	0.40 0.29 0.40 0.30

¹ These values are averages only. The accurate determination of elastic constants is a very difficult problem and much depends on the purity and heat treatment of the material. The values for Aluminum, Copper, Steel, Glass, Lucite, Polystyrene are taken from W. C. Schneider and C. J. Burton, J. App. Phys. 20, 48, 1949. Most of the other values are from "Handbook of Chemistry and Physics" (Chemical Rubber Publishing Co., Cleveland, 30th ed., 1948). All values are given for room temperature and in the case of the liquids for approximately atmospheric pressure.

The positive constant k is called the *bulk modulus*, or *modulus* of volume elasticity. We may perhaps regard it as the most significant of the moduli of elasticity since it applies to fluids as well as to solids. Its reciprocal is called the *compressibility* of the substance. Some characteristic values will be found in the preceding table of elastic moduli. It is important to note that since δ_V is non-dimensional, k has the dimensions of p.

Considering next linear strains, if we have a substance in the form of a wire or rod and the longitudinal or tensile stress applied is f_t , with corresponding linear strain δ_t , the ratio

$$\frac{f_t}{\delta_t} = Y \tag{10.1-20}$$

is constant for many solids over a considerable range of values of f_i , and the constant Y so defined is called Young's modulus. Like k it has the dimensions of force per unit area.

Finally if the shearing stress be denoted by f_s and the corresponding shearing strain is θ , we have

$$\frac{f_s}{\theta} = \mu, \qquad (10 \cdot 1 - 21)$$

where μ is constant over a wide range and is known as the *shear modulus* or *rigidity* of the substance. It has the same units as k and Y. These are the three principal moduli of elasticity for homogeneous, isotropic media. We have now to inquire concerning the possible relations among them.

10.2. Relations Among the Elastic Moduli. Consider a rectangular parallelepiped (Fig. 10.4) of sides a, b, c, in the directions of the x, y, z axes respectively, where a < b < c. Suppose that the tensile stress f_t acts along the z axis on the surface FDEB. Then there will be a stress f_t acting along the negative z axis also on the surface CGAO, since the body as a whole is assumed to be in equilibrium. Associated with f_t there will be the linear strain in the z direction f_t/Y and the concomitant lateral strain $\sigma f_t/Y$ perpendicular to the z axis, where σ is Poisson's ratio. Now suppose on

¹It should be emphasized that we are here confining our attention to homogeneous, isotropic media whose properties are the same in every part and in every direction. This rules out consideration of crystalline media whose elastic properties are in general very complicated.

the other hand that f_t acts along the x axis. We then have a linear strain f_t/Y along the x axis and a concomitant lateral strain $\sigma f_t/Y$ perpendicular to the x axis. Finally suppose that

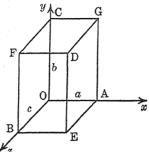


Fig. 10.4

 f_t acts along the y axis; there is then a linear strain f_t/Y along this axis and the concomitant lateral strain of_t/Y perpendicular to the y axis. If now all these stresses act simultaneously we have total linear strains in the x, y, z directions respectively as follows:

$$\delta_x = \delta_y = \delta_z = \frac{f_t}{Y} (1 - 2\sigma).$$
 (10·2-1)

But in this case the resulting deformation is essentially that due to a uniform stress $f_t = p$ on each face of the rod. This is equivalent to a compressive stress and the corresponding volume strain is

$$\delta_V = \frac{f_t}{k} \cdot \tag{10.2-2}$$

Now we have shown (Sec. 10·1) that a volume strain δ_V is equivalent to three mutually perpendicular linear strains each equal in magnitude to $\delta_V/3$. Hence we must have from $(10\cdot2-1)$

$$\frac{f_t}{3k} = \frac{f_t}{Y} (1 - 2\sigma), \qquad (10.2-3)$$

whence there follows the important relation

$$Y = 3k(1 - 2\sigma), (10.2-4)$$

connecting Y, k, and σ .

The shear modulus μ may now be introduced. We have already shown that a shearing strain θ may be replaced by two linear strains $\theta/2$ in magnitude at right angles to each other and of opposite sign. Hence in the present case if the positive linear stress f_t acts along the z axis and the negative stress $-f_t$ acts along the x axis the result will be a shearing strain at an angle of 45° to both linear strains and of magnitude

$$\theta = \frac{2}{Y} f_t(1+\sigma). \tag{10.2-5}$$

Now it can be shown, however, that a tensile stress f_t in the z direction and a lateral negative stress $-f_t$ in the x direction are equivalent to a shearing stress of magnitude f_t acting at an angle of 45° to the other stresses.

For suppose we consider the cross section of the parallelepiped in the xz plane with the stresses f_t as indicated (Fig. 10.5). Cut the parallelepiped with a plane parallel to the y axis and with trace LC in the xz plane making angle 45° with AG. Now the resultant of the force Sf_t acting on the area S parallel to the y axis (the trace of which on the xz plane is GL) and Sf_t acting on the equal area (the trace of which is CG)

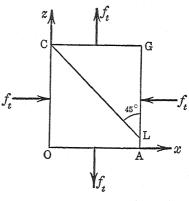


Fig. 10.5

will be in magnitude $\sqrt{2}Sf_t$ and its direction will be parallel to the 45° plane just introduced. Hence the stress on this latter plane will be $\sqrt{2}Sf_t/\sqrt{2}S = f_t$, if we recall that the area of the 45° plane is $\sqrt{2}$ times the areas it cuts off on the faces of the parallelepiped. Hence the statement made above is justified. But if the shearing stress $f_s = f_t$ acts, the resulting shearing strain is

$$\theta = \frac{f_t}{\mu},\tag{10-2-6}$$

where μ is the shear modulus. This must be the same shearing strain as that given in $(10\cdot2-5)$, for its direction is the same and it corresponds to precisely the same set of external stresses acting on the body. Hence we may equate and get

$$\frac{1}{\mu} = \frac{2}{Y} (1 + \sigma), \tag{10.2-7}$$

or

$$Y = 2\mu(1+\sigma),$$
 (10·2–8)

an important relation connecting Y, μ and σ . By means of (10·2-4) and (10·2-8), the bulk and shear moduli k and μ respectively may be expressed in terms of Young's modulus and Poisson's

308 DEFORMABLE BODIES AND WAVE MOTION

ratio. By the elimination of the latter between the two equations we may also obtain the relation connecting Y, k, and μ

$$Y = \frac{9k\mu}{\mu + 3k},\tag{10.2-9}$$

showing that the three moduli are not independent; from a knowledge of any two, the third may be derived. If now we eliminate Y between the eqs. (10·2–4) and (10·2–8), we obtain some information about σ . Thus

$$3k(1-2\sigma) = 2\mu(1+\sigma). \tag{10.2-10}$$

It is to be noted that k and μ are essentially positive quantities. Hence in order that both sides of $(10\cdot2-10)$ may have the same sign, σ must not be greater than $\frac{1}{2}$. On the other hand it must not be less than -1, since otherwise the term $1 + \sigma$ would be negative while $1 - 2\sigma$ would be positive. Hence we may state the important inequality relation for σ

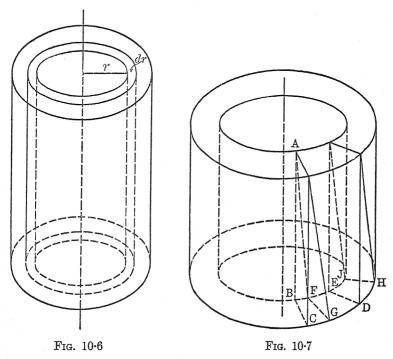
$$-1 < \sigma < \frac{1}{2}.\tag{10.2-11}$$

It is interesting to notice in the table at the end of the preceding section that σ for many substances has values ranging around $\frac{1}{4}$ or $\frac{1}{3}$. Poisson devised a theory of elasticity according to which σ for all elastic solids should be exactly $\frac{1}{4}$. The theory is however not particularly well substantiated except as far as order of magnitude is concerned.

It may be well to mention here briefly some of the experimental methods for determining the elastic moduli. The evaluation of Young's modulus is doubtless familiar to the student as it forms a part of almost every elementary laboratory routine. The length of a long vertical wire of the substance to be examined is measured when stretched by a weight. A known weight is then added and the elongation very carefully measured with a micrometer microscope. This enables the linear strain δ_l to be calculated, and the stress having been computed from the extra weight and the area of cross section of the wire, Y follows from the formula defining it.

The determination of μ or k is a more difficult matter. Usually μ is the quantity investigated, whence k may be calculated by means of eq. (10·2–9). We shall discuss only one method of getting the shear modulus, the theory of which is interesting on its own

account. This is the torsion of a right circular cylinder. Consider (Fig. 10-6) such a solid in the form of a wire of radius a and height h fastened vertically at the upper end. Suppose now that a torque in the form of a couple is applied to the lower end of the wire twisting the wire at this end through an angle θ about its axis. Any cylindrical element or sheet of the wire of radius r



and thickness dr will suffer a shear. The magnitude of the shearing strain will be understood from the following figure (Fig. 10-7) where an enlarged version of a portion of the cylindrical sheet (of unit height, let us say) is presented. Consider a part of this ring enclosed between two diametral planes. This is marked in dotted lines in the figure (A-BCDE) and is approximately a parallelepiped. Now when the wire is twisted and the ring sheared, the part just considered assumes the shape shown by the full lines (A-FGHJ). The shearing strain is then approximately \overline{BF} since \overline{AB} has been chosen of unit length. But $\overline{BF} = r\phi$, where ϕ is the angle of twist per unit height of the element. It follows that since

the twist is θ for the cylinder of height h, the shearing strain for the ring element of radius r but height h will be

$$\frac{r\theta}{h}.$$
 (10·2–12)

Then the shearing stress f_s will necessarily be given as

$$f_s = \frac{\mu r \theta}{h}, \qquad (10 \cdot 2 - 13)$$

where μ is the shear modulus of the material composing the wire. The shearing stress is the force per unit area on the base of the cylindrical element, and hence the *torque* or force moment per unit area about the axis of the cylinder due to the shearing stress is

$$rf_s = \frac{\mu r^2 \theta}{h}, \qquad (10 \cdot 2 - 14)$$

while the torque for the whole base area of the element is

$$2\pi r \, dr \cdot r f_s = \frac{2\pi \mu r^3 \theta \, dr}{h}. \tag{10.2-15}$$

The total torque on the base area of the *complete* cylinder is obtained by integrating the preceding expression from 0 to a. Finally for the shearing torque we have

$$M = \frac{\pi\mu a^4}{2h}\theta. \tag{10.2-16}$$

Suppose now that a heavy disc or rod is rigidly attached to the free end of the wire, and that its moment of inertia about the axis of the wire is I. If the disc is turned through an angle θ from its equilibrium position there will be exerted on it a restoring torque of magnitude M given by eq. (10·2–16). Hence the equation for the rotational motion of the disc after its release will be (see Sec. 7·5)

$$I\ddot{\theta} = -\frac{\pi\mu a^4}{2h}\theta. \tag{10.2-17}$$

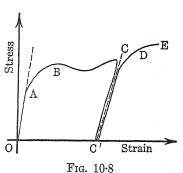
But from our previous study we observe that this is the equation of angular simple harmonic motion, with frequency

$$\nu = \frac{1}{2\pi} \sqrt{\frac{\pi \mu a^4}{2hI}}.$$
 (10·2–18)

The experimental measurement of ν , h, a, and I suffices to determine μ , the rigidity, from the above expression.

10.3. Elastic Limit. Fatigue and Heredity. Our treatment of the first two sections has implied the validity of Hooke's law. For the most interesting applications of the study of deformable bodies this law will generally be assumed, together with the existence of the elastic moduli and their relations. However, it is well to

recognize that this law is valid only within limits. Suppose we take a piece of wire and strain it, plotting the tensile stress against the linear strain. The result in general will be somewhat as shown in Fig. 10.8. From O to A the stress is very closely a linear function of the strain, i.e., Hooke's law holds. As long as we remain within this region of stress and strain the wire behaves



as an elastic medium, i.e., resumes its original length on the removal of the stress. However, after the strain indicated by A has been reached, further stress does not lead to a corresponding strain in accordance with the relation

$$\frac{\text{stress}}{\text{strain}} = K, \tag{10.3-1}$$

where K is a constant. Rather the stress-strain curve becomes bent as in the region from A to B: a small increase in stress leads to a much larger increase in strain than the above relation (10·3–1) would indicate. On release from stress in this state the wire does not recover its original length at once and sometimes not even after a very long time. The point A may be said to mark the elastic limit of the original specimen while the point B represents what may be called the yield point, because beyond it the elongation increases very rapidly and the wire becomes plastic in that the strain depends on the time a given stress acts: if the stress is maintained the wire may and in many cases does break. However, in some cases the specimen appears to regain its elasticity after the point B has been reached and for such the curve begins to mount

again. Suppose in this process, say at point C, the load is removed. It is then found that the original stress-strain curve is not followed at all, but that the change follows a curve like CC', more or less straight, so that the complete removal of stress now leaves a permanent strain, of magnitude OC'. This is known as a permanent set. The further application of stress leads to a new stress-strain curve following C'C and also more or less straight, i.e., effectively the elastic limit has been raised by the stretching and is now beyond the yield point of the unstretched specimen. It should be noted that the interval between the elastic limit and the breaking point varies greatly from substance to substance, being greater for ductile substances as one might expect.

The stress-strain curve in Fig. 10·8 is a static curve, i.e., the values of the strain there shown are the equilibrium values attained after the stress has been applied. We have already commented in Sec. 9·2 on the fact that the strain never builds up to its full value immediately on application of stress, but that time is taken for this process. This served as an illustration of the concept of relaxation time. The temporal change of strain in a solid subjected to stress has been called "creep" by metallurgists and is obviously of considerable practical importance in materials which have wide commercial use. Much attention to this phenomenon has recently been paid by physicists.

Within the elastic limit a wire will return to its original length on release from stress but this process also shows a relaxation effect. Substances differ greatly in the rate at which the original length is regained and with some a considerable time is required. For example while quartz shows almost no delay, that for a glass fiber may amount to several hours, as may readily be shown by a simple experiment (more easily perhaps for torsion than for extension). This is known as the elastic after-effect or lag. Its magnitude appears to depend on the amount of non-homogeneity in the structure of the substance. This is probably why it is so small in quartz, for example, since quartz is very homogeneous. This is true of crystals in general. Glass on the other hand is a composite mixture of fairly large aggregates and hence less homogeneous. It is possible

¹ Cf. for example, F. Seitz, "The Physics of Metals" (McGraw-Hill, New York, 1943). This book discusses in considerable detail both the elastic and plastic properties of solids, particularly metals, both polycrystalline and the single crystal state.

to give a rather good purely mechanical analogy of the after-effect phenomenon by means of the model indicated in Fig. 10-9. A weight B is attached to the horizontal support A by means of the spring S_1 and by means of the spring S_2 to the weight C which is immersed in a very viscous liquid (molasses, for example). Now suppose that B is moved down a little. If it is held for such a short time that C in the meantime has not had a chance to move perceptibly, when B is released it will spring back to its original position (oscillating slightly about it). But if B is kept in the

displaced position long enough, C will move down slowly, and when B is released it will not oscillate about its original position at once, since this would correspond to a stretching of the spring S_2 . After a time C regains original equilibrium position whereupon B does the same. Another thing we should notice is that the oscillations of B die away the more rapidly the greater the viscosity of the liquid in which C is immersed, i.e., the greater the elastic after-effect. We should therefore expect that when wires which show a large aftereffect are set vibrating either longitudinally or torsionally the vibrations will become rapidly damped out. This is a very striking effect and may be thought of as due to a kind of

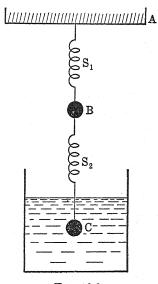


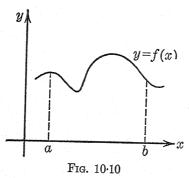
Fig. 10.9

internal viscosity of the metal very similar to viscosity in fluids. There is an interesting property of this viscosity of metals which was discovered by Lord Kelvin. Suppose a wire is forced to vibrate for a considerable length of time; it is then found that the rate of damping for free oscillations is much greater than for the same wire before it was vibrated. The wire acts as if it were tired. On being allowed to rest, recovery ensues and the rate of damping decreases again. This effect is known as elastic fatigue. Another illustration of the same effect is found in the fact that repeated application of a stress may so weaken a metal that it breaks at less than the normal breaking point.

In the phenomena of elasticity we find an interesting illustration of a new physical concept. It has been seen that the ordinary stress-strain relation holds only over a limited range and even here the phenomena of relaxation, elastic after-effect, and fatigue present themselves. It is just as if the elastic state of a wire at a given instant does not depend on the strain at that instant but on the whole previous strain history of the wire. The question then appears to be one of heredity: the inheritance of the previous states is a controlling factor in any present state. The problem is particularly interesting because it introduces a new type of mathematical method into physics. We shall illustrate this briefly. Strictly speaking, even within the elastic limit, we ought to write Hooke's law in the form

$$\delta = Kf + \phi, \tag{10.3-2}$$

where f is the stress, δ the strain, K a constant, and ϕ is a quantity which depends on all the values which f has taken on from the time



when the first stress was applied to the specimen up to the time t being considered. A quantity of this sort which depends on a whole range of values of another quantity is known as a functional. The simplest illustration is provided by the area under a curve between two ordinates. Thus consider the curve y = f(x) (Fig. 10·10). The area included between the curve, the x axis and

the two ordinates at x = a and x = b is a functional of f(x), for it clearly depends on the whole set of values of y = f(x) between a and b. We may write

$$A = \int_{a}^{b} f(x) \ dx \equiv F[[f_{a}^{b}(x)]], \qquad (10.3-3)$$

to indicate this type of quantity.

Let us illustrate briefly by a definite problem, the torsion of a wire. Letting θ be the angle of torsion and M the magnitude of the applied torque, we can rewrite eq. $(10\cdot3-2)$ as follows,

$$\theta = KM + \phi, \tag{10.3-4}$$

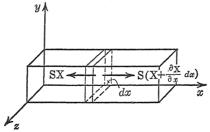
where M is a function of time, viz., M(t). We shall now assume that unit torque applied to the wire during a time interval of $d\tau$ from τ to $\tau + d\tau$ not only produces a certain torsion in the wire at this moment but also contributes a torsion at the future time t of amount $\phi(t, \tau) d\tau$. Hence the resultant torsion angle θ at the time t will be given by KM(t) plus the sum of all the residual terms $\phi(t, \tau)M(\tau) d\tau$. This sum of course must be an integral over all the time elapsing from the initial treatment at time t_0 to t. So we now write for the angle the "generalized" Hooke's law

$$\theta(t) = KM(t) + \int_{t_0}^{t} \phi(t, \tau) M(\tau) d\tau.$$
 (10.3-5)

If we knew the previous torque at every instant and also knew the function $\phi(t,\,\tau)$ we might compute the resultant torsion. The function $\phi(t,\,\tau)$ is known as the coefficient of heredity, and the eq. (10·3–5) is known as an integral equation because in general knowing $\theta(t)$ and $\phi(t,\,\tau)$ we are interested in finding M(t), which occurs under the integral sign. The solution of this problem is too difficult for presentation here. We merely wish to emphasize the fact of its existence in an endeavor to describe analytically the phenomena of elastic after-effect and fatigue, with which it might be supposed to be rather hopeless to deal in any symbolic way. The student interested in hereditary elasticity and in heredity problems in physics in general will find some interesting material in Volterra, "Theory of Functionals" (London, 1930). Integral equations are becoming constantly more important in the solution of physical problems.

10.4. Wave Motion. We now wish to discuss the motion of an elastic solid. This is in general a difficult subject. However, there is a particular type of such motion which is of such importance throughout physics that we must give it considerable attention. This is known as wave motion. Suppose that a strain is produced at some point in an elastic medium. What happens throughout the rest of the medium? For the sake of simplicity (though with no loss in generality as far as the fundamental ideas are concerned), in answering this question we shall confine ourselves here to the special case of an infinitely long wire or rod placed along the x axis and subjected to a tensile stress which we shall denote by X. This is a force per unit area acting solely along the

x axis. Consulting Fig. 10.11, where a portion of the rod is indicated and represented by a rectangular parallelepiped, we



Frg. 10.11

consider the motion of a thin section of the rod of length dx. The area of cross-section being S, the stress force on the element by the portion of the rod to the left is

while that due to the portion of the rod to the right is

$$S\left(X + \frac{\partial X}{\partial x} \, dx\right) \cdot$$

Hence the net force in the positive x direction is $S \frac{\partial X}{\partial x} dx$. The equation of motion results from equating this to the kinetic reaction $\rho S \xi dx$ where ρ is the mean density of the rod material and ξ is the displacement of the section considered to move as a whole. We therefore have

$$\rho \ddot{\xi} = \frac{\partial X}{\partial x} \cdot \tag{10.4-1}$$

Now from Hooke's law there at once results

$$X \div \frac{\partial \xi}{\partial x} = Y, \tag{10.4-2}$$

where Y is Young's modulus [Sec. 10·1, eq. (10·1-20)], and $\frac{\partial \xi}{\partial x}$ is the linear strain. On substitution from (10·4-2) into (10·4-1) we get the equation

$$\rho \ddot{\xi} = Y \frac{\partial^2 \xi}{\partial x^2} \cdot \tag{10.4-3}$$

The analysis leading to $(10\cdot 4-3)$ is faulty in that it neglects the fact that the *size* of the element $S\ dx$ changes as it moves. However, this turns out to be a very small change if ξ is small and hence the equation is correct for *small* displacements.

In order to understand how the various parts of the rod move we must solve to find ξ as a function of x and t. Eq. 10·4-3 is a partial differential equation of the second order and a detailed discussion of its solution would require a good deal of mathematics. We shall however proceed as simply as possible. Since the equation says that the second time derivative of ξ is a constant times the second space derivative, it is natural to inquire whether or not a solution in the general form

$$\xi = f(x + ct) \tag{10.4-4}$$

is possible, where f is an arbitrary differentiable function and c is a constant, whose value is to be determined in such a way as to fit the solution. On this assumption then

$$\dot{\xi} = cf', \quad \ddot{\xi} = c^2f'', \quad (10.4-5)$$

where we have set

$$f' = \frac{df(x+ct)}{d(x+ct)},$$

and correspondingly for f''. Moreover, similarly

$$\frac{\partial^2 \xi}{\partial x^2} = f''. \tag{10.4-6}$$

If now we resubstitute into the differential equation (10.4-3) we get

$$\rho c^2 f^{\prime\prime} = Y f^{\prime\prime}, \qquad (10.4-7)$$

which shows that our choice of a solution works provided we have

$$c = \pm \sqrt{\frac{Y}{\rho}}$$
 (10.4-8)

In other words c may be either positive or negative but must have the magnitude $\sqrt{Y/\rho}$. Since both $+\sqrt{Y/\rho}$ and $-\sqrt{Y/\rho}$ yield solutions, we may have either

$$\xi = f_1(x - ct),$$
 (10.4-9)

or

$$\xi = f_2(x + ct). \tag{10.4-10}$$

Since moreover

$$\frac{\partial^2 (f_1 + f_2)}{\partial t^2} = \ddot{f}_1 + \ddot{f}_2,$$

and

$$\frac{\partial^2 (f_1 + f_2)}{\partial x^2} = \frac{\partial^2 f_1}{\partial x^2} + \frac{\partial^2 f_2}{\partial x^2},$$

it follows that

$$\xi = A_1 f_1(x - ct) + A_2 f_2(x + ct) \tag{10.4-11}$$

is also a solution where A_1 and A_2 are arbitrary constants. In other words the *sum* of two solutions of the differential equation (10·4–3) each multiplied by an arbitrary constant is also a solution. The same is true of *any* linear combination of solutions. This is an important property of *linear* differential equations, i.e., those in which no squares or higher powers of the dependent variable or any of its derivatives enter.

Let us now inquire into the physical significance of the solution (10·4-11). We first consider f_1 alone. The value of the function f_1 at the point $x = x_0$ and the time $t = t_0$ is

$$\xi_{00} = f_1(x_0 - ct_0). \tag{10.4-12}$$

Its value at x_0 at the later time t_1 is

$$\xi_{01} = f_1(x_0 - ct_1),$$
 (10.4–13)

where ξ_{01} is in general different from ξ_{00} . But if we take its value at t_1 at the point x_1 , where

$$x_1 - x_0 = c(t_1 - t_0),$$
 (10.4-14)

we clearly have

$$\xi_{11} = f_1[x_0 + c(t_1 - t_0) - ct_1]$$

$$= f_1(x_0 - ct_0)$$

$$= \xi_{00}.$$
(10.4-15)

In other words the value of ξ for x_0 , t_0 is the same as its value for x_1 , t_1 , provided the distance from x_0 to x_1 is equal to the time elapsing between t_0 and t_1 multiplied by c. It is just as if the value of ξ had been propagated from x_0 to x_1 in the time interval $t_1 - t_0$ with the velocity c. Hence $f_1(x - ct)$ may be taken to represent a disturbance (denoted by ξ) which is propagated in the positive x direction with velocity equal to c. We call this propagated

disturbance a wave and the type of motion described is wave motion. Since ξ is a function of t and x only, the wave here defined is known as a plane wave, corresponding to propagation in one direction only. The matter is rendered more evident from an examination of the figure (Fig. 10·12) where we have plotted two functions of x, namely $f(x-ct_0)$ and $f(x-ct_1)$, which are the values of ξ for the time instants t_0 and t_1 respectively. It is as if we had taken snapshot pictures of the way ξ varies with x at the two instants t_0 and t_1 . The result is the same for both times except that the whole figure is shifted bodily so that each point of the first picture is displaced through the distance $c(t_1-t_0)$ to make the second picture. The

reader is urged to investigate all this very carefully for himself in order to form a clear notion of what a wave really means. He should note carefully, for example, that while to be sure the motion of particles or parts of the medium is involved, the wave motion is not their motion, but rather the motion through the medium

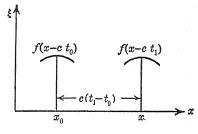


Fig. 10.12

of the configuration which they produce. Thus in the physical problem under discussion what we have shown is that if the long wire is stretched at some place in the direction of its length, thus producing a disturbance of the parts of the wire in the vicinity of this point and then let go, the disturbance moves along the wire with velocity $c = \sqrt{Y/\rho}$, constituting what is called an elastic wave. Moreover the wave travels in both directions, for if $f_1(x-ct)$ corresponds to wave motion in the positive x direction, $f_2(x+ct)$ will correspond to similar motion with the same velocity in the negative x direction. It must be emphasized, moreover, that there is nothing essentially periodic about this motion, though, as we shall see, periodic or harmonic waves are undoubtedly the most important type. We ought also to note that in the case discussed the direction of propagation coincides with the direction of displacement: the wave is said to be longitudinal (see Sec. 10.5).

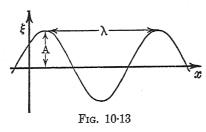
The numerical magnitude of the velocity of an elastic wave in a rod or wire is of some interest. If we consider steel, for example, with $\rho = 7.8 \text{ grams/cm}^3$ and $Y = 2 \times 10^{12} \text{ dynes/cm}^2$ (varying

of course with the special composition), we get $c=5.06\times 10^5$ cm/sec = 5060 meters/sec. This indeed is the velocity with which sound travels along a steel bar, since the sound is propagated by elastic disturbance.

It is now in order to discuss in a general way some of the properties of *simple harmonic* waves. For a wave of this kind progressing in the positive x direction the displacement may be written

$$\xi = A \cos \kappa (x - ct), \qquad (10.4-16)$$

where A is called the *amplitude* of the wave, c is the velocity and κ is a constant which must have the dimensions of a reciprocal distance, since the argument of the cosine must be non-dimensional



(i.e., a pure number). It is seen that (10.4-16) actually has the form of the function f_1 in (10.4-11). Now let us plot ξ as a function of x for a certain instant of time t. The result is given in Fig. 10.13. In wave nomenclature the places where ξ is a maximum

are called *crests*; those corresponding to minima are called *troughs*. The distance between any two successive crests is called one *wave length* and designated by λ . Since the cosine function has a period of 2π , i.e.,

$$\cos z = \cos (2\pi + z),$$

it follows that for any instant t

$$\kappa(x-ct) + 2\pi = \kappa(x+\lambda-ct), \qquad (10.4-17)$$

and therefore

$$\kappa = \frac{2\pi}{\lambda} \cdot \tag{10.4-18}$$

Moreover at any particular place there is simple harmonic motion [recall that $\cos (-\kappa ct) = \cos \kappa ct$] of frequency

$$\nu = \frac{\kappa c}{2\pi} = \frac{\omega}{2\pi} \,, \tag{10.4-19}$$

so that

$$\kappa = \frac{\omega}{c} = \frac{2\pi\nu}{c} \cdot \tag{10.4-20}$$

We shall call ν the frequency of the wave, but shall often use $\omega = 2\pi\nu$ in its place for convenience. Finally the period of the simple harmonic wave is

$$P = \frac{2\pi}{\omega} = \frac{1}{\nu}, (10.4-21)$$

and we have the important relation for a simple harmonic wave

$$P = \frac{2\pi}{\kappa c} = \frac{\lambda}{c}, \qquad (10.4-22)$$

the period appearing as the ratio of wave length to velocity. It is clear that we may write the expression for ξ in the following equivalent forms

$$\xi = A \cos \frac{2\pi}{\lambda} (x - ct)$$

$$= A \cos (\kappa x - \omega t)$$

$$= A \cos 2\pi \left(\frac{x}{\lambda} - \frac{t}{P}\right) \cdot (10.4-23)$$

The choice is a matter of convenience. We shall perhaps use

$$\xi = A \cos(\omega t - \kappa x), \qquad (10.4-24)$$

most frequently, reversing the order of the time and space terms. This is, of course, purely arbitrary.

At this place it may be desirable to point out the value of the complex notation in expressing displacement in wave motion. Thus analogously to eq. (9.4-5) in Sec. 9.4 we set

$$\xi = Ae^{i(\omega t - \kappa x)}, \qquad (10.4-25)$$

where the amplitude A is now in general to be considered complex. In operating with wave displacements, particularly in differentiation and integration, this form of expression is very advantageous. Of course it should be emphasized that in general only the real or the imaginary parts separately have physical significance.

The wave dealt with in the last few paragraphs has been one traveling in the positive x direction. Naturally all that has been said will apply equally well to the wave in the negative x direction, where the displacement is denoted by

$$\xi = Be^{i(\omega t + \kappa x)}. \tag{10.4-26}$$

In most cases where solid media are in question there will be waves traveling in both directions simultaneously and the resultant displacement will then be the sum of (10·4–25) and (10·4–26). We shall discuss special cases in Sec. 10·5.

The quantity $\omega t \pm \kappa x$ is called the *phase* of the wave, provided the amplitude is real. If the latter is complex, we can always write it $A = A_0 e^{i\epsilon}$, where A_0 and ϵ are real quantities. The phase now takes the form $\omega t \pm \kappa x + \epsilon$, and ϵ is called the initial phase at the origin.

When a medium is traversed by a wave, a certain amount of energy is associated with the motion in each unit volume. In the case of a harmonic wave the kinetic energy may be easily calculated. Suppose such a wave travels in the x direction along a rod of cross-sectional area S. Then the kinetic energy in a length Δx of the rod comprising a volume $S \Delta x$ is

$$\frac{1}{2}\rho S\dot{\xi}^2 \Delta x, \qquad (10.4-27)$$

for ρS Δx is the mass involved, if ρ is the density of the rod. Hence the kinetic energy per unit volume at any place x at a given time instant t is $K = \frac{1}{2}\rho \dot{\xi}^2$. Now for $\dot{\xi}$ we have, consulting eq. (10·4–24) and supposing the initial phase zero,

$$\dot{\xi} = -A\omega \sin(\omega t - \kappa x), \qquad (10.4-28)$$

so that the kinetic energy per unit volume or kinetic energy density becomes

$$K = \frac{1}{2}\rho A^2 \omega^2 \sin^2(\omega t - \kappa x).$$
 (10.4-29)

This is of course a function of x and t. At any place it varies periodically with time and at any instant it varies periodically from point to point. Rather more important is the *average* kinetic energy density. Thus for any value of x we have for the time average

$$\frac{t}{K} = \frac{1}{2}\rho A^2 \omega^2 / \tau \cdot \int_0^{\tau} \sin^2(\omega t - \kappa x) dt, \qquad (10.4-30)$$

where τ is much greater than one period of the wave. Since x is considered constant, the evaluation of the integral gives the same result as (9·3–13). Thus

$$\vec{K} = \frac{1}{2}\rho A^2 \omega^2.$$
 (10.4–31)

Keeping the time constant, let us find the *space* average of the kinetic energy density over a distance long compared with the wave length, viz.,

$$\bar{K} = \frac{1}{2}\rho A^2 \omega^2 / x_0 \cdot \int_0^{x_0} \sin^2(\omega t - \kappa x) \, dx, \qquad (10.4-32)$$

where $x_0 \gg \lambda$, the wave length. The form of the integral is precisely the same in the two cases and hence

$$\vec{\overline{K}} = \vec{K}. \tag{10.4-33}$$

Therefore we may speak merely of the average kinetic energy density as

$$\overline{K} = \frac{1}{4}\rho A^2 \omega^2. \tag{10.4-34}$$

Since the medium is strained by the passage of the wave there will also be potential energy associated with the wave. This will, however, obviously depend on the character of the strain corresponding to the wave and hence must be specially investigated in each particular case. In this place as a simple illustration let us consider again the solid rod of density ρ and Young's modulus Y. The potential energy associated with the element Δx of the rod is the work done by the variable force due to stress, viz., SX, in changing the length Δx from $(\Delta x)_1$ to $(\Delta x)_2$, i.e., by $\Delta(\Delta x)$. Thus we have

$$V = S \int_{(\Delta x)_1}^{(\Delta x)_2} X d(\Delta x). \tag{10.4-35}$$

Now the change $d(\Delta x)$ is the differential of the linear strain times the original length. Hence we may write

$$d(\Delta x) = d\left(\frac{\partial \xi}{\partial x}\right) \Delta x. \tag{10.4-36}$$

Consequently (10·4-35) becomes

$$V = S\Delta x \int_0^{\frac{\partial \xi}{\partial x}} X d\left(\frac{\partial \xi}{\partial x}\right)$$
$$= YS \Delta x \int_0^{\frac{\partial \xi}{\partial x}} \frac{\partial \xi}{\partial x} d\left(\frac{\partial \xi}{\partial x}\right), \qquad (10.4-37)$$

using (10.4-2). We finally obtain

$$V = \frac{YS \,\Delta x}{2} \left(\frac{\partial \xi}{\partial x}\right)^2 \cdot \tag{10.4-38}$$

For the plane harmonic wave (10·4-24) we get for the potential energy density

$$V = \frac{YA^{2}\kappa^{2}}{2}\sin^{2}(\omega t - \kappa x). \qquad (10.4-39)$$

It is of interest to observe from a comparison with (10·4-29) that for this type of wave motion

$$V = K, \qquad (10.4-40)$$

identically. Hence the total energy density is

$$U = \rho \omega^2 A^2 \sin^2 (\omega t - \kappa x). \qquad (10.4-41)$$

The average total energy density is therefore

$$\bar{U} = \frac{1}{2}\rho\omega^2 A^2. \tag{10.4-42}$$

We must be careful not to conclude that the relation (10·4–40) is true for all kinds of waves. Nevertheless the special case here discussed is rather important in practice.

Another important allied concept which deserves mention in this place is that of the *intensity* of a wave. This is defined as the average rate of flow of energy per unit area perpendicular to the direction of propagation. If we consider any cross-section of the rod, for example, as the wave passes this cross-section it may be thought of as carrying energy with it and the amount of energy carried per unit area per second is clearly (for a plane wave at least) the product of the energy density and the velocity of the wave. Hence we have the following expression for the intensity in the case of a plane harmonic wave traversing a rod [from eq. (10·4–42)]

$$I = \frac{1}{2}\rho cA^2\omega^2. \tag{10.4-43}$$

It may perhaps be worth while to express this in terms of the maximum stress in the rod, X_{max} , associated with the wave. Since

$$X = Y \frac{\partial \xi}{\partial x},$$

and

$$\frac{\partial \xi}{\partial x} = A\kappa \sin (\omega t - \kappa x),$$

we have

$$X_{\text{max}} = A\kappa Y$$

$$= A\frac{\omega}{c}\rho c^2 = A\omega\rho c. \qquad (10.4-44)$$

Therefore

$$A^{2}\omega^{2} = \frac{X^{2}_{\text{max}}}{\rho^{2}c^{2}},$$

$$I = \frac{1}{2}\frac{X^{2}_{\text{max}}}{cc}.$$

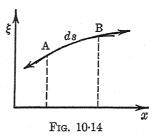
$$(10.4-45)$$

and

The intensity is thus proportional to the square of the maximum tensile stress produced in the rod. This is a very significant equation.

10.5. Transverse Waves in a String. In our study of wave motion in a solid rod in the previous section we were concerned with waves in which the direction of propagation coincides with the material displacement. Such waves are known as *longitudinal* waves and are of great importance, particularly in acoustics (Sec. 11.6). However, there is another type of wave, of equal significance, in which the displacement is perpendicular to the

direction of propagation. Such waves are known as transverse waves. As a matter of fact both types of waves are set up in every disturbance of an elastic solid, though in a fluid only the former can exist. The mathematical discussion of transverse wave motion in a general elastic solid is rather involved and we shall not enter upon



it here. However, we encounter a very interesting illustration of transverse waves in the problem of the vibrating string. Imagine a perfectly flexible string (Fig. 10·14) stretched with tension τ . Let the equilibrium position of the string be the x axis (we neglect its thickness, of course), and denote the *transverse* displacement of any point by ξ . We suppose that all displacements

take place in the same plane. We shall also assume that the tension τ is constant and may be considered to a sufficient degree of approximation as unchanged when the string is deformed. This in turn implies that the displacement ξ is to be very small compared with the length of the string. We now discuss the motion of an element of the string from A to B of length ds, whose projection on the x axis is dx. Let us find the restoring force acting on the element AB in the direction of the displacement. The upward component of τ at A is

$$-\tau \frac{\partial \xi}{\partial s} = -\tau \left[\frac{\partial \xi}{\partial x} - \frac{1}{2} \left(\frac{\partial \xi}{\partial x} \right)^3 + \cdots \right], \qquad (10.5-1)$$

which we get by expanding

$$\frac{\partial \xi}{\partial s} = \frac{\frac{\partial \xi}{\partial x}}{\sqrt{1 + \left(\frac{\partial \xi}{\partial x}\right)^2}},$$
 (10.5-2)

recalling that both $\partial \xi/\partial s$ and $\partial \xi/\partial x$ are small. As a matter of fact, it will be a sufficiently good approximation, if we write for the upward component at A

$$-\tau \frac{\partial \xi}{\partial x} \cdot \tag{10.5-3}$$

What is now the upward component of the tension at B? It must be, to a first approximation,

$$\tau \left[\frac{\partial \xi}{\partial x} + \frac{\partial}{\partial x} \left(\frac{\partial \xi}{\partial x} \right) dx \right], \tag{10.5-4}$$

whence the total upward force on the element AB is the sum of (10.5-3) and (10.5-4) namely

$$\tau \frac{\partial}{\partial x} \left(\frac{\partial \xi}{\partial x} \right) dx. \tag{10.5-5}$$

Equating the kinetic reaction of the element to the force on it yields the equation of motion

$$\rho \ddot{\xi} \, dx = \tau \, \frac{\partial^2 \xi}{\partial x^2} \, dx,$$

or

$$\rho \ddot{\xi} = \tau \frac{\partial^2 \xi}{\partial x^2}, \qquad (10.5-6)$$

where ρ is here the mass per unit length of the string and assumed to be a constant. We have here made the same sort of approximation as in deriving (10·4-3) by neglecting the stretching of the string element in its displacement. Eq. (10·5-6) is the usual wave equation, and hence the displacement ξ is propagated along the string with velocity

$$c = \sqrt{\frac{\tau}{\rho}}$$
 (10.5–7)

If we recall the discussion in Sec. 10.4, we see that since in general there may be waves in both directions the most general expression for the displacement at distance x and at time t is

$$\xi = f_1(ct - x) + f_2(ct + x). \tag{10.5-8}$$

Let us, however, confine our attention to the motion of the string when the displacement at any point is simple harmonic, i.e., consider only harmonic waves. From Sec. 10.4 we then write

$$\xi = Ae^{i(\omega t - \kappa x)} + Be^{i(\omega t + \kappa x)}, \qquad (10.5-9)$$

where A and B are complex amplitudes, $\omega = 2\pi\nu$, ν being the frequency of the wave and $\kappa = \omega/c$, where c is the velocity [eq. $(10\cdot5-7)$]. There are now four arbitrary constants involved, since A and B, being complex, each involve two constants. Moreover there is nothing to specify the frequency, which is therefore also arbitrary.

We shall now suppose that the string is finite and of length l, and is moreover fastened at the ends so that no motion takes place there. We then have the boundary conditions that for all t

$$\xi = 0 \text{ for } x = 0, l.$$
 (10.5–10)

Substitution into (10.5-9) then yields

$$(A + B)e^{i\omega t} = 0,$$

 $e^{i\omega t}(Ae^{-i\kappa l} + Be^{i\kappa l}) = 0.$ (10.5–11)

From the first of these two equations we get

$$A = -B, (10.5-12)$$

and hence from the second

$$e^{-i\kappa l} - e^{i\kappa l} = 0. (10.5-13)$$

Expressing eq. (10·5–13) in terms of trigonometric functions gives

$$\sin \kappa l = 0, \qquad (10.5-14)$$

which leads to

$$\kappa l = n\pi, \qquad (10.5-15) \bullet$$

where n is any integer. This immediately limits the possible frequencies of the harmonic waves in the string to the set given by

$$\nu_n = \frac{nc}{2l} = \frac{n}{2l} \sqrt{\frac{\tau}{\rho}}$$
 (10.5–16)

The lowest frequency of the set, viz., $\nu_1 = c/2l$, is called the fundamental. The higher frequencies $\nu_2, \ldots, \nu_n, \ldots$ are called the harmonics. It is interesting to see that the imposition of the boundary conditions (10.5–10) in addition to removing some of the arbitrary constants from the solution has also introduced a certain discreteness into the problem. This is a significant result, for it is the first element of discreteness we have encountered in the motion of a continuous medium. Corresponding to each frequency ν_n there is a definite mode of oscillation, viz.,

$$\xi_n = A_n [e^{i(2\pi\nu_n t - \kappa_n x)} - e^{i(2\pi\nu_n t + \kappa_n x)}]. \qquad (10.5-17)$$

For practical purposes we shall put this into its real form. Letting $2A_n = a_n + ib_n$,

$$\xi_{n, \text{ real}} = \frac{a_n}{2} \left[\cos \left(2\pi \nu_n t - \kappa_n x \right) - \cos \left(2\pi \nu_n t + \kappa_n x \right) \right] \\ - \frac{b_n}{2} \left[\sin \left(2\pi \nu_n t - \kappa_n x \right) - \sin \left(2\pi \nu_n t + \kappa_n x \right) \right]. \quad (10.5-18)$$

If we expand the trigonometric expressions and simplify, there finally results

$$\xi_{n, \text{ real}} = \sin \kappa_n x [a_n \sin 2\pi \nu_n t + b_n \cos 2\pi \nu_n t], \quad (10.5-19)$$

with $\kappa_n = 2\pi\nu_n/c$. For any integral value of n, eq. (10.5–19) represents a solution of the wave equation (10.5–6) and a possible displacement propagation for the string. Now the wave equation (10.5–6) is a *linear* partial differential equation, and we have

already seen (Sec. 10·4) that the individual solutions of such an equation are additive, giving new solutions. It would therefore follow that the most general expression for the displacement when harmonic waves of all frequencies ν_n travel along the string, is given by

$$\xi_{\text{real}} = \sum_{n=0}^{\infty} \sin \kappa_n x [a_n \sin 2\pi \nu_n t + b_n \cos 2\pi \nu_n t]. \quad (10.5-20)$$

What does this solution mean physically? It means that if we bring about an initial transverse displacement of any point on the string the resulting displacement in time of all other points is given by the infinite series (10.5–20).

It must be emphasized that (10.5-20) no longer represents a single progressive wave. Rather as the superposition of progressive harmonic waves in opposite directions along the finite string it represents what is called a standing or stationary wave. Thus each point of the string is executing harmonic motion of frequency given by the bracketed terms of (10.5-20) and with amplitudes proportional to $\sin \kappa_n x$ (not a function of time). This introduces a feature not present in the usual progressive wave. For we see that for any particular mode of oscillation characterized by the integer n there exist n-1 equally spaced points of the string (exclusive of the end points) at which the displacement is always zero. These are given by the condition

$$\sin \kappa_n x = 0,$$

and measured from one end of the string the corresponding values of x are

$$x = l/n$$
, $2l/n$, $\cdots (n-1)l/n$.

The points in question are called *nodes* or nodal points. The reader should show that the distance between successive nodes is one half the wave length of the corresponding standing wave. A knowledge of this and the frequency of excitation suffices (from 10·4–22) to permit the evaluation of the velocity of the wave. The midpoints between successive nodes are referred to as loops since there the displacement reaches its maximum during each period of the motion.

Our next task is to discuss the evaluation of the constants a_n and b_n . Let us denote the initial displacement of any point on

the string by ξ_0 . This is a function of x, of course. We therefore have from (10.5-20), (dropping the subscript "real," since this is now understood)

$$\xi_0 = \sum_{n=0}^{\infty} b_n \sin \kappa_n x. \qquad (10.5-21)$$

Moreover, let us denote the initial *velocity* of any point on the string by $\dot{\xi}_0$, also a function of x. Since

$$\dot{\xi} = \sum_{n=0}^{\infty} \sin \kappa_n x \cdot 2\pi \nu_n [a_n \cos 2\pi \nu_n t - b_n \sin 2\pi \nu_n t], \quad (10.5-22)$$

we have

$$\dot{\xi}_0 = \sum_{n=0}^{\infty} 2\pi \nu_n a_n \sin \kappa_n x. \qquad (10.5-23)$$

We now have to evaluate a_n and b_n from the eqs. (10.5-21) and (10.5-23). This might appear to be rather difficult. Let us, however, multiply both sides of eq. (10.5-21) by $\sin \kappa_s x$, where s is a particular integer, and integrate the result as x goes from 0 to l. This gives

$$\int_0^l \xi_0 \sin \kappa_s x \, dx = \sum_{n=0}^\infty \int_0^l b_n \sin \kappa_n x \cdot \sin \kappa_s x \, dx. \quad (10.5-24)$$

From (11.5-16)

$$\kappa_s = \frac{2\pi\nu_s}{c} = \frac{s\pi}{l} \,, \tag{10.5-25}$$

and

$$\int_{0}^{l} \sin \frac{s\pi x}{l} \cdot \sin \frac{n\pi x}{l} dx$$

$$= \left(\frac{\sin (s-n)\pi x/l}{2(s-n)\pi/l} - \frac{\sin (s+n)\pi x/l}{2(s+n)\pi/l} \right) \Big]_{0}^{l}, \quad (10.5-26)$$

so that if $n \neq s$, the integral vanishes. On the other hand if n = s, we have

$$\int_0^l \sin^2 s\pi \, \frac{x}{l} \cdot dx = \frac{l}{2s\pi} \left(\frac{s\pi}{l} x - \sin \frac{s\pi x}{l} \cos \frac{s\pi x}{l} \right) \Big]_0^l = \frac{l}{2} \cdot$$

Hence of all the integrals in the sum on the right-hand side of

(10.5-24) only one is different from zero, namely that for which n=s. Its value is l/2. We therefore have

$$b_s = \frac{2}{l} \int_0^l \xi_0 \sin \kappa_s x \, dx, \qquad (10.5-27)$$

which enables us to calculate b_s if we know ξ_0 . A precisely similar procedure with eq. (10.5–23) leads to

$$a_s = \frac{2}{s\pi c} \int_0^l \dot{\xi}_0 \sin \kappa_s x \, dx. \qquad (10.5-28)$$

What the two eqs. (10.5-27, 28) really mean is this: if we know the actual displacement and velocity of every point on the string at a single definite instant we can compute the configuration of the string, i.e., the position and velocity of every point, for every future instant, tracing out its whole subsequent history. is an interesting bit of mathematics connected with this result. The initial position of the string can be anything compatible with the conditions imposed, i.e., it can be mathematically speaking any continuous function of x. What we have then shown is that any such function of x can be expanded in a certain bounded interval into an infinite series of circular functions of x, viz., (10.5-21), and we have shown how to calculate the coefficients. Such a series is called a Fourier Series, and is of very great importance in many physical problems such as are encountered for example in the theory of the conduction of heat, oscillations of media, orbits in celestial mechanics and atomic structure theory. The student at this point should obtain a reasonable degree of familiarity with this type of series. We have not discussed it here with any rigor, since we have not taken up, for example, the conditions for its convergence. But we have at any rate pointed out some of its physical significance.

Further discussion of the motion of a stretched string lies really in the province of acoustics.²

10.6. Types of Elastic Waves in Solids. We have discussed in the last two sections the special cases of longitudinal waves in a

¹ See E. B. Wilson, "Advanced Calculus" (Ginn & Co., Boston, 1911), p. 458 ff

² See H. Lamb, "Dynamical Theory of Sound" (Macmillan, 2nd ed., New York, 1925), Chap. II.

solid rod and transverse waves in a string. As has already been indicated, the problem of the general motion of an elastic solid disturbed from equilibrium demands rather more extensive analysis than we wish to embark on here. This is particularly true of solids like crystals which have different properties in different directions. However, we ought to note the principal results for the case of an isotropic solid, viz., one whose properties are independent of direction. Such a solid possesses, as we have seen, both shear and volume elasticity. The former involves distortion of the medium and hence implies the existence of transverse waves, while the latter provides for the existence of longitudinal waves. Thus in the general motion of an elastic solid we shall expect to find both types of waves present. The mathematical analysis substantiates this expectation and shows that the longitudinal wave in an extended solid medium travels with a velocity

$$c_L = \sqrt{\frac{k + \frac{4}{3}\mu}{\rho}}, \qquad (10.6-1)$$

while the transverse wave has a velocity

$$c_T = \sqrt{\frac{\mu}{\rho}}, \qquad (10.6-2)$$

where k and μ have their usual significance as elastic moduli (Sec. 10·1). We therefore have $c_L > c_T$ for all solids. It will be recalled that for a long narrow rod the velocity of the longitudinal wave reduces to

$$c = \sqrt{\frac{Y}{\rho}}, \qquad (10.6-3)$$

where Y is Young's modulus. Comparison shows that for a large number of solids c_L is greater than c by about 10%.

The foregoing formulas suggest that the elastic constants of solids may be determined by the measurement of the velocities of the corresponding types of waves. This has indeed been carried out with success by employing standing waves in solid rods similar to those already discussed for the finite string in Sec. 10.5. In recent years it has been possible to drive such rods at very high frequencies (e.g., 100 megacycles). This is particularly true of piezo-electric crystals which alternately contract and expand when

placed across an alternating high voltage. The fact that the modes are very sharply defined by the dimensions of the specimen (cf. the length l in Sec. 10·5) makes it possible to maintain oscillations of very specific frequency and indeed to stabilize the frequency of high frequency electric (radio) circuits in which the vibrating crystals are placed.

10.7. The Elastic Medium Theory of Light. The student will recall from his study of elementary physics that there have been two principal theories for the propagation of light. The first, championed by Newton, is the so-called corpuscular theory according to which light travels as very small particles moving in straight lines through space in all directions from a given source. The second is the wave theory suggested by Huyghens, which considers the transmission of light as a wave propagation in a medium. Of the two theories the latter has been the more successful if we leave out of consideration the phenomena connected with the emission and absorption of light, which seem to demand a particle description based on the ideas of the quantum theory.

Now in our discussion of wave motion we have seen that in every case it is a disturbance in a medium which is propagated as Hence if the transmission of light is a wave propagation, what is the medium which is disturbed? This is a problem to which physicists gave considerable attention during the nineteenth century, beginning with Fresnel, who was the first to put the wave theory on its feet mathematically speaking, and continuing with the work of Green, Neumann, MacCullagh, Lord Kelvin and Lord Rayleigh, all men of outstanding eminence. Let us consider briefly the nature of the problem as an illustration of the mechanical theory developed in this chapter. It was only natural to assume that the light-bearing medium is analogous to an elastic medium of some kind. Now we have noted that in an elastic solid any disturbance is in general propagated by both transverse and longitudinal waves. The only elastic medium through which waves of but one type are propagated is a fluid, as we shall see in the next chapter (for a fluid has volume elasticity only and is unable to support shearing stress). In this case longitudinal waves only are transmitted. The fact that a light wave can be polarized shows definitely that it must be a transverse wave. Hence the light-bearing medium cannot be a fluid, but must be an elastic

solid. Here, however, the problem at once arises: how shall one get rid of the longitudinal wave that accompanies the transverse wave? We have seen that the longitudinal wave travels with

velocity $c_L = \sqrt{\frac{k + 4\mu/3}{\rho}}$, while the transverse wave has the

velocity $c_T = \sqrt{\mu/\rho}$. Therefore it is impossible to get rid of the former type of wave merely by imagining the medium to possess no volume elasticity, viz., k = 0. To overcome this difficulty Green made the assumption that $k = \infty$, so that c_L is infinite. Effectively then the longitudinal wave is not propagated at all, and this would seem to be satisfactory enough. Unfortunately difficulties arise when the theory is applied to the specific problem of the reflection and refraction of light and extended to the case of anisotropic media like crystals. Hence it would seem that Green's theory must be abandoned. Lord Kelvin proposed a theory in which $k = -4\mu/3$ so that $c_L = 0$. Such a medium has negative volume elasticity and hence will be unstable, i.e., a positive stress will lead to expansion instead of contraction. Of course, one might imagine such a medium as filling the whole universe or as being rigidly attached to a fixed containing vessel acting as its boundary. It is hardly necessary to say that the mathematical development of any one of these theories is extremely complicated. Nevertheless Lord Kelvin worked all his life in the endeavor to construct such a mechanical theory of light. It must be confessed that the result has not been particularly successful. We are now inclined to believe that the electromagnetic theory of Maxwell is more satisfactory. Incidentally, however, it should be pointed out that Maxwell originally based his theory on mechanical grounds, so that we are still justified in considering the generally accepted theory of light propagation as one which in the last analysis can be visualized in a dynamical framework.

PROBLEMS

- 1. The shear modulus of a certain substance is equal to 1.24×10^{11} dynes/cm², and Young's modulus for the same substance is 3.2×10^{11} dynes/cm². Calculate the value of the bulk modulus for this substance. Also compute Poisson's ratio. Do the same for a quartz fiber for which $Y = 5.179 \times 10^{11}$ dynes/cm² and $\mu = 2.88 \times 10^{11}$ dynes/cm². Comment on any interesting feature of the result.
- 2. The value of Young's modulus for steel is 2×10^{12} dynes/cm², and the coefficient of linear (thermal) expansion for the same substance is 1.2×10^{-5}

per degree centigrade. What compressive force must be applied to the ends of a steel cylinder 2 cm in diameter to prevent it from expanding longitudinally when the temperature is raised by 40°C?

- 3. A body suffers tensile strains $\delta_x, \delta_y, \delta_z$, along the x, y, and z directions respectively. Show that the resultant strain is equivalent to a uniform volume dilatation equal in magnitude to $(\delta_x + \delta_y + \delta_z)$ plus two shears, one in the xy plane of magnitude $\frac{2}{3}(\delta_x + \delta_y + \delta_z) 2\delta_y$ and the other in the xz plane of magnitude $\frac{2}{3}(\delta_x + \delta_y + \delta_z) 2\delta_z$.
- 4. A steel wire 100 cm long and 0.1 cm in radius is suspended from a rigid ceiling. A homogeneous horizontal disc of mass 1000 grams and radius 10 cm is attached at its center to the free end. How much twisting force must be applied at the periphery of the disc to turn the latter through an angle of 10°? If the disc is released and allowed to move freely, what will be the period of its motion?
- 5. A circular cylindrical wire spring is said to be flat if, when the spring is stretched, each turn of wire is approximately parallel to a plane perpendicular to the axis of the cylinder. One end of such a spring of radius a is attached to a rigid ceiling while the other is stretched by the imposition of a vertical force F. Show that the equivalent stiffness of the spring is given by the expression (when $r \ll a$)

$$f=\frac{\mu\pi r^4}{2a^2l},$$

where $\mu=$ shear modulus of the wire, r= radius of the wire, and l= total length of the wire in the spring. Discuss the significance of the condition $r\ll a$.

- 6. Calculate the potential energy per unit volume in a solid rod subjected to a longitudinal stress of magnitude X.
- 7. A homogeneous solid beam of square cross-section with side a is bent into the arc of a circle of radius R. If x denotes distance measured normally to the central axis of the beam and L is the magnitude of the bending torque at any point about an axis normal to the plane of bending and passing through the central axis, show that the ratio of change of L with x is given by

$$\frac{dL}{dx} = \frac{2Yax^2}{R},$$

where Y is Young's modulus for the material of the beam. Hence find that the total bending torque necessary is

$$L = Ya^4/12R$$

8. A possible dynamical variant of Hooke's law to take account of the "relaxation" effect (see 9.2) may be written in the form

$$X = A\delta + B\dot{\delta},$$

where X is the stress and δ the corresponding strain, while A and B are elastic

constants. Suppose X = 0 for $0 < t < t_0$, $X = X_0 = \text{constant for } t_0 < t < t_1$ and X = 0 for $t > t_1$. Find the way in which δ varies with the time (choosing arbitrary but reasonable values for A and B). Discuss elastic after-effect on this basis.

- 9. Show that $\xi = f(x + ct)$ represents a wave in the negative x direction with velocity c. Investigate by graphical methods or otherwise the physical character of the wave f(x ct) f(x + ct). Do the same for the expression f(x ct) + f(x + ct).
- 10. A wave $\xi = f(x ct)$ progresses in the positive x direction with velocity c. At the point $x = x_0$ a rigid barrier is interposed. Describe the character of the motion of the medium to the left of the barrier.
- 11. Use the relations among μ , k, and Y to compare c_L , c_T , and $c = \sqrt{Y/p}$ (Sec. 10-6). Discuss the physical significance of the difference between c_L and c.
- 12. A plane sinusoidal acoustic wave in air of frequency 512 cycles/sec has a displacement amplitude 10^{-6} cm. If it progresses along the x axis with the velocity of sound at 20° C, and the origin is taken at the point where the displacement ξ is zero at t=0, plot to scale the wave form at t=1 sec for a distance of a wave length or so. Also plot the variation of ξ at x=5 cm as a function of t for two or three periods.
- 13. In the preceding problem, find the displacement velocity, i.e., $d\xi/dt$ at x=5 cm and t=1 sec. Find the acceleration, i.e., $d^2\xi/dt^2$ under the same conditions. Plot both quantities in the same way in which you plotted ξ in Problem 12.
- 14. Show that the intensity of a compressional wave in a rod may be represented by $\overline{X_x} \dot{\xi}$, where the bar indicates the time average. Compare this expression with (10·4-45).
- 15. A flexible string 80 cm long with a mass of 4 grams is stretched with a tension of 84×10^6 dynes. Find the fundamental and first two harmonics of the string. The string is pulled aside at its midpoint a distance 0.5 cm and let go. Find the amplitudes of the fundamental and first two harmonics in the resulting standing wave pattern. Prove that all harmonics of even order are absent from the motion of the string.
- 16. The string in Problem 15 is struck at its midpoint (while in its equilibrium position) with such force as to give it an initial velocity there of 10 cm/sec. Find the amplitudes of the fundamental and first two harmonics in the resulting standing wave pattern.
- 17. Calculate the average kinetic and potential energies for the first two harmonics in the motions of the string discussed in Problems 15 and 16.
- 18. Show that the differential equation for wave motion through a dissipative medium may to a first approximation be written in the form

$$\ddot{\xi} + R\dot{\xi} = c^2 \frac{\partial^2 \xi}{\partial x^2} \cdot$$

Obtain a particular solution of this equation in the form of a damped harmonic wave. Obtain the expression for the distance which a disturbance will travel before its amplitude is diminished in the ratio 1/e. Suggest possible physical characteristics which make a medium dissipative.

- 19. A narrow solid rod of length l and line density ρ is clamped rigidly at its two ends. What are the natural frequencies for longitudinal vibration of the rod? Obtain also the expression for the natural frequencies for the case of a rod clamped at one end and free at the other. Work the same problem for a rod free at both ends but clamped at the center.
- 20. Two harmonic waves of the same amplitude but slightly different frequencies ν and $\nu + \Delta \nu$ move in the x direction with velocities c and $c + \Delta c$ respectively. It is assumed that c depends on the frequency (dispersion) so that $\Delta c = \frac{dc}{d\nu} \Delta \nu$. Show that if the two waves are superposed the resultant may be considered a harmonic wave of frequency ν but with an amplitude which varies with time with frequency $\Delta \nu/2$. Show that the varying amplitude travels with velocity U such that

$$\frac{1}{U} = \frac{d\left(\nu/c\right)}{d\nu} \cdot$$

U is called the group velocity of the compound wave system.



CHAPTER XI

MECHANICS OF FLUIDS

11.1. Fluids at Rest. Fundamental Principles of Hydrostatics. In Chapter X we introduced the idea of a deformable body and discussed its behavior. We saw there that the most general kind of deformation may be considered to be a combination of a dilatation (change of volume) with a shear (change of shape). In the case of physical solids the application of stress yields in general both kinds of strains, but there is a class of substances of such a nature that the application of stress gives rise to dilatation only and never to shears. Such substances are called fluids or better perfect fluids, for they represent limiting ideal cases of the actual fluids encountered in nature. The line between actual solids and actual fluids is sometimes hard to draw. Under very great stresses most solids at ordinary room temperature begin to act like fluids. and a solid like pitch, for example, flows at room temperature under the stress produced by its own weight, though for impulsive stresses (i.e., those of short duration) it may behave like other solids in manifesting both shears and dilatations. Naturally all

Fig. 11·1

substances become fluids at sufficiently high temperatures and appropriate pressures.

From what we have just said it should be possible to deduce the principles governing the behavior of fluids from the fundamental relations for deformable bodies in general. The rigorous derivation demands the use of more general stress-strain

analysis than we have given in the previous chapter. Fortunately we are able to visualize the deduction in a rather simple way. Let us imagine a volume element of fluid as given in the figure (Fig. 11·1), and consider further any plane surface element in

this box. Suppose that the fluid is at rest, i.e., in equilibrium under the action of external forces. There will still be a stress on the surface element; and this stress will be a compressive stress, i.e., one tending to produce dilatation only, since by definition a perfect fluid can not be sheared. Now if the compressive stress on the area (to which we shall hereafter refer as the pressure) were to be inclined to the area at any other angle than 90°, it would have a non-vanishing component along the surface. But this would lead to shearing and, since a perfect fluid can not support a shear, motion would have to ensue parallel to the surface. This contradicts the initial assumption of equilibrium. Hence we reach the important conclusion, amply verified by simple experiments, that in a perfect fluid at rest the stress or pressure on any surface element whatever is normal to the surface. To state the same result in slightly different form, at any point in a perfect fluid the pressure is normal

to any surface element passing through this point and independent of the direction of this element. This statement includes the famous principle of Pascal.

Let us now examine the condition that a fluid may be in equilibrium under the action of an external force F acting on unit mass. Let F_x , F_y , F_z denote the x, y, z components of F respectively. Consider an

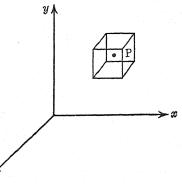


Fig. 11.2

element of fluid contained in the parallelepiped with sides dx, dy, and dz (Fig. 11·2). Let the pressure at the midpoint P of the volume be p. Then the normal force on the face parallel to and nearer the yz plane is $\left(p - \frac{\partial p}{\partial x} \frac{dx}{2}\right) dy dz$, while that on the farther parallel face is $\left(p + \frac{\partial p}{\partial x} \frac{dx}{2}\right) dy dz$. Hence the fluid in the element is acted on by a force due to pressure whose component in the x direction is

$$-\frac{\partial p}{\partial x} dx dy dz. (11\cdot1-1)$$

The analogous components in the y and z directions are $-\frac{\partial p}{\partial y}$

dx dy dz and $-\frac{\partial p}{\partial z} dx dy dz$. Now if the element is to be in equi-

librium under the influence of the external force, the total force components on the element in the x, y, and z directions must vanish. That is, we must have (recalling that F represents the force per unit mass)

$$\rho F_x dx dy dz - \frac{\partial p}{\partial x} dx dy dz = 0,$$

$$\rho F_y dx dy dz - \frac{\partial p}{\partial y} dx dy dz = 0,$$

$$\rho F_z dx dy dz - \frac{\partial p}{\partial z} dx dy dz = 0,$$

$$(11.1-2)$$

or more compactly,

$$F_x = \frac{1}{\rho} \frac{\partial p}{\partial x}$$
, $F_y = \frac{1}{\rho} \frac{\partial p}{\partial y}$, $F_z = \frac{1}{\rho} \frac{\partial p}{\partial z}$, (11·1-3)

where ρ is the density. These are the fundamental equations of hydrostatics. We see that if we write the change in pressure between two closely neighboring points in the fluid at rest as dp, we have the total differential

$$dp = \rho F_x dx + \rho F_y dy + \rho F_z dz. \qquad (11.1-4)$$

The problem of ascertaining the distribution of pressure in the fluid is the problem of solving this total differential equation. We may express the general solution in the form

$$p = \Phi(x, y, z) + C,$$
 (11·1-5)

where C is a constant of integration whose value depends on the pressure at some specified point in the fluid. If in some way the pressure were to be increased at this point, eq. $(11\cdot1-5)$ states that the pressure is increased by the same amount at every other point in the fluid. This is sometimes referred to as the law of the transmissibility of pressure. It is usually considered a part of Pascal's principle for fluids at rest.

We shall consider one special case, namely that in which the

external force is gravity. Suppose that it acts downward along the z axis, i.e.,

$$F_x = F_y = 0, F_z = -g.$$
 (11·1-6)

Then p varies with z alone and we have

$$\frac{dp}{dz^2} = -\rho g. \tag{11.1-7}$$

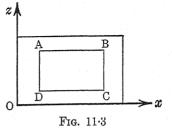
The integration can be carried out only if we know the way in which ρ depends on z. If the fluid is incompressible (i.e., an ideal liquid) ρ is constant, and we have

$$p = -\rho gz + C, \tag{11.1-8}$$

C being an arbitrary constant which may be put equal to zero if the origin is taken at a point where the pressure is zero, viz., at the surface of the liquid (assuming that there is a vacuum above the surface). It is to be noted that the negative sign enters (11·1-8) because the positive direction of z is upward and p decreases as one goes up. In practical applications it is usually more convenient to measure z downward and use the positive sign. (See, for example, Sec. 11·2.) In words, (11·1-8) states that in an incompressible fluid at rest under the action of gravity alone the pressure varies directly as the depth. This is well substantiated by experimental investigation of liquids in which the variation of ρ with depth is negligible for moderate depths.

11.2. Principle of Archimedes — Stability of Floating Bodies. We can apply the results of the preceding section very appro-

priately to the problem of the calculation of the resultant force exerted by a liquid on an object immersed in it. Let us first take a special case for the sake of simplicity. Imagine a rectangular parallelepiped with dimensions a, b, c immersed so that one pair of faces is parallel to the surface of the liquid. In Fig. 11·3



the cross-section is shown, where $\overline{AB} = a$, and $\overline{BC} = c$. Suppose that \overline{AB} is at a distance z_0 below the surface. To calculate the total force exerted by the liquid on the parallelepiped we must find the x,

y, and z components. The x component of the force, for example, will be found by multiplying the element of area at every point by the x component of the pressure at that point and then summing up over the whole surface of the body. The procedure for the other components is analogous. Now from the way we have chosen our special case it is clear that the x and y components will be zero since the forces on opposite faces cancel out in pairs. The z component, however, is not zero. Let us suppose that the pressure at the surface of the liquid is p_0 . At the top of the body (\overline{AB}) it is then by eq. $(11\cdot1-8)$

$$p_1 = p_0 + \rho g z_0,$$

while at the bottom (\overline{DC}) it is

$$p_2 = p_0 + \rho g(z_0 + c).$$

The areas of the top and bottom are the same, viz., S = ab, so that there is a *downward* force on the top face of magnitude

$$S(p_0 + \rho g z_0),$$

and an upward force on the bottom face of magnitude

$$S(p_0 + \rho g(z_0 + c)).$$

Hence the total force in the z direction is an upward force of magnitude

$$F_z = \rho g c S. \tag{11.2-1}$$

Now since cS is the volume of the body and pgcS accordingly the weight of the liquid displaced by the body, it follows that the liquid exerts on the immersed body a buoyant force equal to the weight of the displaced liquid. This is the principle of Archimedes. We have derived it here for a very special case. In the general case of a body of any shape we should have for the component of force in the z direction (gravity alone being assumed to act)

$$F_z = \iint p \cos \gamma \, dS, \qquad (11.2-2)$$

where p is the pressure acting on the area element dS and γ is the angle between the normal to dS and the z axis. The integration is to be carried out over the *whole* surface of the body. Since gravity

acts vertically along the z axis we see that even in the general case the x and y components are zero, i.e.,

$$F_x = F_y = 0. (11.2-3)$$

In order to calculate F_z it is convenient to transform the surface integral into a volume integral. Suppose we consider again an elementary parallelepiped dx, dy, dz cut out of the body as shown in Fig. 11-2. We have already seen in the preceding section that the net force on this element in the z direction is

$$-\frac{\partial p}{\partial z} dx dy dz. (11\cdot 2-4)$$

Consequently we can get the resultant force on the whole body by evaluating the integral

$$-\iiint \frac{\partial p}{\partial z} dx dy dz, \qquad (11.2-5)$$

the integration being extended over the whole volume of the body; this must also give us F_z . In other words the volume integral $(11\cdot2-5)$ is equivalent to the surface integral $(11\cdot2-2)$. We have here a special simplified case of the so-called divergence theorem or, more generally, Green's theorem, which the student will encounter in all advanced work in physics, and of which we shall give a more extensive account in the next section. Now we know from

eq. (11·1-7) that in the present case $\frac{\partial p}{\partial z} = -\rho g$, where ρ is the density of the liquid. Hence for the resultant force we get

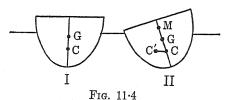
$$F_z = \rho g \iiint dx \, dy \, dz$$
$$= \rho g V, \qquad (11.2-6)$$

where V is the volume of the body. The result is then the same as for the special case. It should be noted that although in our discussion we have used the word "liquid," everything we have said is true of *fluids* in general.

We have thus seen that a liquid exerts a buoyant force on an object immersed in it. This force is directed vertically upward and is in magnitude equal to the weight of the displaced liquid. We can also see that it must pass through the center of gravity of the body; for if we imagine the body to be removed, the force

which the rest of the liquid exerts on the portion of liquid taking the place of the body must remain the same. But if this force did not pass through the center of gravity this particular portion of the fluid would not be in equilibrium with regard to rotation, and hence the hydrostatic conditions of the problem would be violated.

Let us next give our attention to the special case of a body floating on the surface of a liquid. In this case the buoyant force must clearly be numerically the same as the weight of the body. Hence by the principle of Archimedes a floating body displaces its own weight of liquid. We have just seen that the buoyant force acts upward through the center of gravity of the displaced liquid.



This is called the *center of buoyancy* of the body. The weight of course acts through the center of gravity of the body itself. Hence a floating body is acted on by a *couple*, whose moment vanishes only when the center of gravity and the center of buoyancy lie in the same vertical line.

If a floating body is slightly displaced from its position of equilibrium, while the center of gravity G of the body (I and II, Fig. 11-4) remains fixed in the body, the center of buoyancy C will necessarily move with respect to it, since the shape of the immersed portion has now altered. Suppose that it moves from C to C' (II, Fig. 11-4), tracing out the curve CC' in the body. The center of curvature of this curve, M, is called the metacenter and its height above the center of gravity is called the metacentric height. We shall denote it by h. For not too great angles of roll, Mmay be considered as the point of intersection of the vertical through C' and the line \overline{CG} extended. The metacentric height is the important criterion for the stability of a floating object, e.g., a ship. If this height is positive, i.e., if M lies above G, the rolling produces a couple which tends to right the ship; while, if it is negative $(M \ below \ G)$ the resulting couple tends to produce further rolling. The equilibrium is therefore stable in the former case and unstable in the latter. If M coincides with G, the equilibrium will

be neutral, which is just as bad as instability as far as a ship is concerned. It is clear that the higher the metacentric height the greater the stability. However, it can be shown that the period of the roll varies as the inverse square root of the metacentric height. Hence a ship with a large h, generally known as a "stiff" ship, will roll more rapidly and hence not be so comfortable as a slowly rolling ship for which h is smaller. The latter is known as a "crank" ship. The tendency in ocean ships is to build them with as small a metacentric height as is consistent with safety.

11.3. The Equation of Continuity in Fluid Motion. As in the discussion of the motion of the vibrating string (Sec. 10.5), we shall carry through our analysis of fluid motion by concentrating attention on a small element of volume and observing its behavior over a short interval of time. We shall find that there are two fundamental ideas involved. The first is the expression of the fact that the body in question acts like a continuous medium and not merely as a discrete aggregate of particles.² We must consider once more the elementary parallelepiped dx dy dz (Fig. 11.2) with the point P as its midpoint. Assume that this element is stationary and that the perfect fluid whose motion is being investigated flows through it. We shall suppose that the flow of the fluid is such that at time t, the components of the velocity of flow at P(x, y, z) are u, v, w in the three coördinate directions respectively. The x component of the velocity of flow at the face parallel to the yz plane and nearer to it will then be $u - \frac{\partial u}{\partial x} \frac{dx}{2}$

while that on the farther parallel face will be $u + \frac{\partial u}{\partial x} \frac{dx}{2}$. Hence the rate of flow of fluid, i.e., the mass per second into the one face and out of the other will be

$$\left(\rho u - \frac{\partial (\rho u)}{\partial x} \frac{dx}{2}\right) dy dz, \qquad (11\cdot 3-1)$$

and

$$\left(\rho u + \frac{\partial (\rho u)}{\partial x} \frac{dx}{2}\right) dy dz, \qquad (11.3-2)$$

¹ See A. G. Webster, "Dynamics," p. 474.

² The second, the idea involved in the equations of motion, will be treated in the next section.

respectively, if the density of the fluid at P at time t is ρ . Therefore the excess of *outflow* over *inflow* will be given by the difference of $(11\cdot3-2)$ and $(11\cdot3-1)$ or

Excess flow in
$$x$$
 direction = $\frac{\partial (\rho u)}{\partial x} dx dy dz$. (11·3-3)

Similarly we find for the excess flow in the y and z directions, $\frac{\partial (\rho v)}{\partial y} dx dy dz$ and $\frac{\partial (\rho w)}{\partial z} dx dy dz$ respectively. Hence the total excess outflow is

$$\left(\frac{\partial (\rho u)}{\partial x} + \frac{\partial (\rho v)}{\partial y} + \frac{\partial (\rho w)}{\partial z}\right) dx dy dz.$$
 (11·3-4)

But since the fluid is a continuous medium and can neither be created nor destroyed the excess outflow must be compensated by a decrease in the mass of the fluid contained within the element. If the density at time t is ρ , the rate of change is $\partial \rho / \partial t$, and hence the rate of decrease of mass in the element is

$$-\frac{\partial \rho}{\partial t} dx dy dz. \tag{11.3-5}$$

Equating (11·3-4) and (11·3-5) we have the equation

$$\dot{\rho} + \frac{\partial(\rho u)}{\partial x} + \frac{\partial(\rho v)}{\partial y} + \frac{\partial(\rho w)}{\partial z} = 0.$$
 (11·3-6)

This is known as the equation of continuity, one of the most important equations in the physics of continuous media. It may be worth while to point out that this equation holds for any deformable continuous indestructible medium, though it plays perhaps its most significant role in connection with fluids. If the fluid is incompressible and homogeneous, the equation of continuity reduces to the simpler form

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0. (11 \cdot 3 - 7)$$

¹ Note that, strictly speaking, the velocity is not $u - \frac{\partial u}{\partial x} \frac{dx}{2}$ or $u + \frac{\partial u}{\partial x} \frac{dx}{2}$ over the *whole* face in each case, as it may vary with y and z. However, our assumption is equivalent to neglecting differentials of order higher than the first. Or, if the student wishes, he may equally well consider the velocities in question as *averages* over the faces concerned.

We can understand the physical significance of this equation more clearly if we consider the flow of fluid through any closed surface of arbitrary form. It is clear that if we divide the surface into a large number of area elements dS, we can get the total rate of flow (in mass per second) out of the surface by multiplying the density by the component of the velocity normal to each area element and further by dS (in each case the outward normal is to be considered positive), and then integrating over the whole surface. Denoting the magnitude of the normal velocity (a function of x, y, z, t) by q_N , we have for the total outflow

$$\iint \rho \ q_N \ dS. \tag{11.3-8}$$

But we have already seen that we can represent the total outflow through an elementary parallelepiped by (11·3-4), and hence the flow through the whole portion of the fluid contained inside the surface considered is given by the volume integral

$$\iiint \left[\frac{\partial (\rho u)}{\partial x} + \frac{\partial (\rho v)}{\partial y} + \frac{\partial (\rho w)}{\partial z} \right] dx \, dy \, dz. \tag{11.3-9}$$

Now (11·3-8) and (11·3-9) represent the same thing and therefore may be equated. If the normal to the area element dS has the direction cosines $\cos \alpha$, $\cos \beta$, $\cos \gamma$ we can at once write

$$q_N = u \cos \alpha + v \cos \beta + w \cos \gamma, \qquad (11.3-10)$$

and on the equating of (11.3-8) and (11.3-9) we finally have

$$\iint \rho (u \cos \alpha + v \cos \beta + w \cos \gamma) dS$$

$$= \iiint \left[\frac{\partial (\rho u)}{\partial x} + \frac{\partial (\rho v)}{\partial y} + \frac{\partial (\rho w)}{\partial z} \right] dx dy dz. \quad (11\cdot3-11)$$

This is a form of the *divergence theorem* or Green's theorem already mentioned in Sec. 11·2. If we look upon u, v, and w as the components of a vector \mathbf{q} , the velocity of the fluid, then

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z}$$

is called the divergence of q, written div q. Similarly

$$\frac{\partial (\rho u)}{\partial x} + \frac{\partial (\rho v)}{\partial y} + \frac{\partial (\rho w)}{\partial z} = \operatorname{div} \rho \mathbf{q}. \tag{11.3-12}$$

There is an interesting consequence of the vanishing of the divergence of $\rho \mathbf{q}$, which as we have just seen [eq. (11·3–7)] is the case for a homogeneous incompressible fluid. At every point in space the velocity \mathbf{q} has a definite magnitude and direction. We can then construct curves whose tangents at every point are in the direction of \mathbf{q} at that point. These curves are called lines of flow. Let us suppose that \mathbf{q} does not change with the time so that the lines

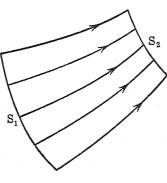
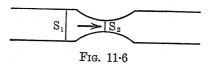


Fig. 11.5

maintain a definite permanent shape. They are then called stream lines. It is clear that a group of them may be considered as generating a surface; this is called a tube of flow (see Fig. 11.5). Now suppose we have such a tube of flow and terminate it by the surfaces S_1 and S_2 respectively. Assume that the fluid is incompressible so that the divergence is zero. The total mass of fluid flowing per second out of the element of volume formed by the

bounded tube is therefore likewise zero. But since there can be no flow either out or in through the sides of the tube (which are parallel to the lines of flow), it must follow that the total normal flow in through S_1 equals the total normal flow out through S_2 . This may be applied at once to the flow of a liquid through a tube or pipe with a constriction (Fig. 11.6). Letting the cross-sectional



area at the unconstricted part be S_1 and that of the constriction be S_2 , if we consider the flow through a portion of the tube with boundaries S_1 and S_2 , it immediately follows from the above that

$$S_1 u_1 = S_2 u_2, (11.3-13)$$

where u_1 and u_2 are the velocities at S_1 and S_2 respectively. For the flow through a constricted pipe the velocity at any point

varies inversely as the area of cross-section. We shall have occasion to note a practical application of this in Sec. 11.5.

11.4. The Equations of Motion of a Perfect Fluid. In order to discuss further the motion of a perfect fluid we must consider its acceleration. Suppose at the point P(x, y, z) the velocity components are u, v, w at the time t, while at the nearby point Q(x + dx, y + dy, z + dz) the corresponding quantities at the time t + dt are u + du, v + dv, w + dw. Since du, dv, dw depend on x, y, z, and t, we have

$$du = \frac{\partial u}{\partial t} dt + \frac{\partial u}{\partial x} dx + \frac{\partial u}{\partial y} dy + \frac{\partial u}{\partial z} dz,$$

$$dv = \frac{\partial v}{\partial t} dt + \frac{\partial v}{\partial x} dx + \frac{\partial v}{\partial y} dy + \frac{\partial v}{\partial z} dz,$$

$$dw = \frac{\partial w}{\partial t} dt + \frac{\partial w}{\partial x} dx + \frac{\partial w}{\partial y} dy + \frac{\partial w}{\partial z} dz.$$

$$(11.4-1)$$

Now if the particle of fluid which was at P at time t is at Q at time t + dt, it follows that

$$dx = u dt$$
, $dy = v dt$, $dz = w dt$,

and du, dv, dw will be the increments in the velocity components of this particle in the time dt. Hence

$$\frac{du}{dt} = \frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} + w \frac{\partial u}{\partial z},$$

$$\frac{dv}{dt} = \frac{\partial v}{\partial t} + u \frac{\partial v}{\partial x} + v \frac{\partial v}{\partial y} + w \frac{\partial v}{\partial z},$$

$$\frac{dw}{dt} = \frac{\partial w}{\partial t} + u \frac{\partial w}{\partial x} + v \frac{\partial w}{\partial y} + w \frac{\partial w}{\partial z},$$
(11-4-2)

will be the components of the acceleration of the particle which was at P at time t and is at Q at time t+dt. The reader must distinguish carefully between $\frac{du}{dt}$ and $\frac{\partial u}{\partial t}$. The latter refers merely to the rate of change of u with time at a particular place. The former gives the genuine rate of change of u for a particle moving from place to place.

Now let us concentrate our attention once more on the element of volume shown in Fig. 11-2. Suppose that at any instant the pressure at P, the center of the parallelepiped, is p. We have already seen that the components of the force on the fluid in the $\frac{\partial p}{\partial p}$, $\frac{\partial p}{\partial p}$, $\frac{\partial p}{\partial p}$

box due to pressure are $-\frac{\partial p}{\partial x}dx\,dy\,dz$, $-\frac{\partial p}{\partial y}dx\,dy\,dz$, $-\frac{\partial p}{\partial z}dx\,dy\,dz$

in the x, y, and z directions respectively. If no external forces in addition to the pressure act on the fluid we may obtain the equations of motion by equating the kinetic reaction components for a unit volume to the corresponding force components. Thus we have, denoting the density once more by ρ , and dividing through by $dx \, dy \, dz$,

$$\frac{du}{dt} = -\frac{1}{\rho} \frac{\partial p}{\partial x}, \quad \frac{dv}{dt} = -\frac{1}{\rho} \frac{\partial p}{\partial y}, \quad \frac{dw}{dt} = -\frac{1}{\rho} \frac{\partial p}{\partial z}. \quad (11.4-3)$$

On the other hand, let us suppose that in addition to the force due to pressure an external force F with components F_x , F_y , F_z acts on unit mass of the fluid. The equations then become

$$\frac{du}{dt} = F_z - \frac{1}{\rho} \frac{\partial p}{\partial x}, \quad \frac{dv}{dt} = F_y - \frac{1}{\rho} \frac{\partial p}{\partial y}, \quad \frac{dw}{dt} = F_z - \frac{1}{\rho} \frac{\partial p}{\partial z}. \quad (11.4-4)$$

A first integration of these equations for the case of an incompressible fluid yields interesting information about the energy of the fluid. Let us assume that the external force is associated with a potential Ω , i.e., such that (see Sec. 4·1)

$$F_x = -\frac{\partial \Omega}{\partial x}$$
, $F_y = -\frac{\partial \Omega}{\partial y}$, $F_z = -\frac{\partial \Omega}{\partial z}$. (11.4-5)

Let us multiply through the eqs. (11.4-4) by u, v, w respectively and add. We obtain

$$\begin{split} u\,\frac{du}{dt} + v\,\frac{dv}{dt} + w\,\frac{dw}{dt} \\ &= -\left(u\,\frac{\partial\Omega}{\partial x} + v\,\frac{\partial\Omega}{\partial y} + w\,\frac{\partial\Omega}{\partial z}\right) - \frac{1}{\rho}\left(u\,\frac{\partial p}{\partial x} + v\,\frac{\partial p}{\partial y} + w\,\frac{\partial p}{\partial z}\right). \end{split} \tag{11-4-6}$$

¹ These are the hydrodynamic equations of motion in the form due to Euler. It may be noted that if, instead of concentrating attention on a given element of space and discussing the flow through this element as time passes, we study the whole history of every particle of fluid as it moves through space, we obtain the equations of motion in the form due to Lagrange. For most purposes those of Euler are more valuable. For those of Lagrange see Horace Lamb, "Hydrodynamics" (Cambridge University Press, 5th ed., 1924), p. 12 ff.

Multiplying through by ρ , recalling that $u = \frac{dx}{dt}$, etc., and noting that $\frac{d(u^2)}{dt} = 2u\frac{du}{dt}$, we obtain

$$\frac{1}{2} \rho \frac{d}{dt} (u^2 + v^2 + w^2) + \rho \frac{d\Omega}{dt} = -\left(u \frac{\partial p}{\partial x} + v \frac{\partial p}{\partial y} + w \frac{\partial p}{\partial z}\right) \cdot (11 \cdot 4 - 7)$$

If, assuming that the fluid is homogeneous and incompressible, we multiply $(11\cdot4-7)$ by the volume element $dx\ dy\ dz$ and integrate over the whole space occupied by the fluid, we have

$$\begin{split} \frac{d}{dt} \left\{ \iiint \frac{1}{2} \rho (u^2 + v^2 + w^2) \, dx \, dy \, dz + \iiint \rho \Omega \, dx \, dy \, dz \right\} \\ &= - \iiint \left(u \, \frac{\partial p}{\partial x} + v \, \frac{\partial p}{\partial y} + w \, \frac{\partial p}{\partial z} \right) dx \, dy \, dz. \quad (11.4-8) \end{split}$$

Now $u^2 + v^2 + w^2$ is the square of the velocity of the fluid and hence $\frac{1}{2}\rho(u^2 + v^2 + w^2)$ is the kinetic energy per unit of volume. Therefore the integral of the latter over the whole space occupied by the fluid must be the total kinetic energy of the fluid, which we may denote by K. The second integral, viz., that of $\rho\Omega$, must represent the total potential energy of the fluid due to the external forces, since Ω is the potential energy for unit mass in so far as these forces are concerned, and $\rho\Omega$ is the corresponding energy per unit volume. We shall denote the second integral by V_{Ω} . The last integral may be transformed into a surface integral over the surface bounding the volume of the fluid, by means of eq. (11·3–11). Let us substitute ρ for ρ in that equation; we have

$$\iiint p \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} \right) dx \, dy \, dz$$

$$+ \iiint \left(u \, \frac{\partial p}{\partial x} + v \, \frac{\partial p}{\partial y} + w \, \frac{\partial p}{\partial z} \right) dx \, dy \, dz$$

$$= \iint p (u \cos \alpha + v \cos \beta + w \cos \gamma) \, dS. \quad (11.4-9)$$

Since the fluid is incompressible, we can write

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0 (11.4-10)$$

from the equation of continuity. Hence resubstituting into eq. $(11\cdot4-8)$ yields

$$\frac{d}{dt}(K+V_{\Omega}) = -\iint p(u\cos\alpha + v\cos\beta + w\cos\gamma) dS. \quad (11\cdot 4-11)$$

The physical significance of the right-hand side may be seen when we recall that $u\cos\alpha + v\cos\beta + w\cos\gamma = q_N$, i.e., is the velocity of the fluid normal to the surface element dS. But $p\,dS$ is the force on dS due to the pressure, and the product of force and velocity is the rate at which work is done by the force. Hence eq. (11.4-11) states that the time rate of change of the total energy of the fluid is equal to the rate at which work is being done by the pressure forces on the boundary surface of the fluid, or alternatively to the rate at which energy flows across this surface.

This equation takes on a particularly interesting form in the case where the pressure p at any place does not vary with the time, i.e., $\partial p/\partial t = 0$. We see that we can now write

$$u\frac{\partial p}{\partial x} + v\frac{\partial p}{\partial y} + w\frac{\partial p}{\partial z} = \frac{dp}{dt}.$$
 (11·4-12)

Therefore if we look upon the quantity $\iiint p \, dx \, dy \, dz$ as the potential energy of the incompressible fluid due to the pressure, we may write $(11\cdot4-8)$ in the form

$$\frac{d}{dt}(T + V_{\Omega} + V_{P}) = 0, \qquad (11.4-13)$$

where $V_{\Omega} + V_{p} = V$ may now be taken as the *total* potential energy of the fluid. In this case (11·4–13) expresses the fact that the total energy is constant, i.e., the fluid now acts like a *conservative system*.

11.5. Steady Flow of a Liquid. Bernoulli's Theorem and Applications. One of the simplest cases of fluid motion is that in which the velocity at any place does not change with the time, even though it changes from place to place. This means that

$$\frac{\partial u}{\partial t} = \frac{\partial v}{\partial t} = \frac{\partial w}{\partial t} = 0. \tag{11.5-1}$$

Such motion is known as steady flow and the lines of flow (Sec.

11.3) are stream lines, as has already been indicated. Let us denote the elementary distance along such a stream line by ds. The direction cosines of the line tangent to the stream line at any point will be $\frac{dx}{ds}$, $\frac{dy}{ds}$, $\frac{dz}{ds}$ respectively, and the component velocities of the fluid, if $q = \sqrt{u^2 + v^2 + w^2}$ denotes the resultant velocity, will be

$$u = q \frac{dx}{ds}$$
, $v = q \frac{dy}{ds}$, $w = q \frac{dz}{ds}$ (11.5–2)

Hence for steady flow the equations of motion (11.4-4) become

$$\frac{du}{dt} = q \left(\frac{\partial u}{\partial x} \frac{dx}{ds} + \frac{\partial u}{\partial y} \frac{dy}{ds} + \frac{\partial u}{\partial z} \frac{dz}{ds} \right) = F_x - \frac{1}{\rho} \frac{\partial p}{\partial x}, \quad (11.5-3)$$

with the two similar equations for y and z. They may be simplified at once to

$$q\frac{du}{ds} = F_x - \frac{1}{\rho} \frac{\partial p}{\partial x}, \quad q\frac{dv}{ds} = F_y - \frac{1}{\rho} \frac{\partial p}{\partial y}, \quad q\frac{dw}{ds} = F_z - \frac{1}{\rho} \frac{\partial p}{\partial z}, \tag{11.5-4}$$

where $\frac{du}{ds}$, etc., denote the gradients (i.e., space rates of increase) of the velocity components along the stream line. A very significant result is obtained by multiplying the equations through by the direction cosines $\frac{dx}{ds}$, $\frac{dy}{ds}$, and $\frac{dz}{ds}$ respectively and adding. We get

$$u\frac{du}{ds} + v\frac{dv}{ds} + w\frac{dw}{ds} = -\frac{d\Omega}{ds} - \frac{1}{\rho}\frac{dp}{ds}, \qquad (11.5-5)$$

the assumption again being made that there exists a potential Ω for the force F. If now we multiply through by ds and integrate along the stream line we obtain

or

$$\int_{s} (u \, du + v \, dv + w \, dw) = -\int_{s} d\Omega - \int_{s} \frac{dp}{\rho},$$

$$\frac{1}{2}q^{2} = -\Omega - \int \frac{dp}{\rho} + C, \qquad (11.5-6)$$

where C is a constant of integration. The $\int \frac{dp}{\rho}$ can not be evaluated until we know the relation between p and ρ . But if we are dealing with a liquid (an incompressible fluid, effectively), ρ is constant and the equation becomes

$$\frac{1}{2}\rho q^2 + \rho\Omega + p = C'. \tag{11.5-7}$$

This equation is an expression for the famous theorem of Bernoulli. It states in words: the total energy per unit volume along any stream line in steady flow is constant, though of course the value of the constant will in general change from one stream line to another. We note that in the above expression $\frac{1}{2}\rho q^2$ is the kinetic energy per unit volume along the stream line, while $\rho\Omega$ and p appear as the potential energy per unit volume due to the external force and the pressure respectively. It is indeed possible to establish Bernoulli's theorem from the independent assumption of the conservation of energy. (See Problem 12 at the end of the chapter.)

Bernoulli's theorem is of considerable importance in practical hydraulics and we shall discuss a few applications here. Suppose in the first place we consider the steady flow through a horizontal tube with a constriction (Fig. 11-6). If we follow a single stream line from the place where the cross-sectional area is S_1 to the place where it is S_2 , the velocity will change from q_1 to q_2 and the pressure from p_1 to p_2 while the potential energy due to the external force, which in this case is gravity, will change very little since there is little or no change in level involved. From the theorem we have accordingly

$$\frac{1}{2}\rho q_1^2 + p_1 = \frac{1}{2}\rho q_2^2 + p_2. \tag{11.5-8}$$

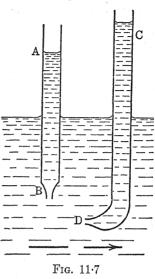
It therefore follows that at the place in the tube where the velocity is greatest the pressure is least and vice versa. We have already seen (Sec. 11·3) that in the steady flow through a tube the velocity varies inversely as the cross-section. Hence where the cross-section is smallest the pressure is least. An interesting practical application of this is the *Venturi water meter*. Substituting $q_2 = \frac{S_1}{S_2}q_1$ into (11·5–8) and solving for q_1 gives

$$q_1 = S_2 \sqrt{\frac{2(p_2 - p_1)}{\rho(S_2^2 - S_1^2)}},$$
 (11.5-9)

whence a knowledge of the two areas of cross-section and of the difference in pressure at the two places, such as might be obtained

by a suitably connected manometer, will give the velocity and also the discharge rate (volume flow per second) of the liquid in the tube.

A rather simple device for measuring fluid velocity is the Pitot tube, illustrated schematically in Fig. 11·7. Two tubes AB and CD are placed with narrowed ends B and D in the pipe through which the fluid is flowing. The tube AB is strictly normal to the flow while CD is bent so that the orifice D faces the flow. We should then expect approximately, at any rate, that the pressure of the liquid inside AB will be p, the same as that in the flowing liquid at this place. On the other hand,



when the steady state is attained the pressure inside the tube CD will be p', where

$$p' = p + \frac{1}{2}\rho q^2, \tag{11.5-10}$$

from Bernoulli's theorem, noting that the velocity in CD is zero. We then have simply

$$q = \sqrt{\frac{2}{\rho} (p' - p)}, \qquad (11.5-11)$$

the velocity being given at once in terms of the difference in pressure in the two tubes as measured by the difference in height of the liquid in the tubes. If this difference is h we have (Sec. 11.1)

$$p' - p = \rho ah, \qquad (11.5-12)$$

and the velocity becomes

$$q = \sqrt{2gh}. (11.5-13)$$

It must be emphasized that the application of the theorem of Bernoulli to this case is a rather bold approximation since the condition of the liquid in the neighborhood of the two openings is hardly compatible with the existence of steady flow. However, eqs. (11.5-11) and (11.5-13) are empirically justified to a fair degree of approximation.

Another application of the Bernoulli theorem is provided in the flow of a liquid out of an orifice in the bottom of a tank. Consider the tank in Fig. 11.8 with the orifice B and suppose that the liquid in the tank is maintained at the level A by a steady fresh supply. If we follow a given stream line from A to B, we have for the

energy per unit volume in the stream line at A,

$$\frac{1}{2}\rho q_A^2 + \rho g h_A + p_A, \quad (11.5-14)$$

while that at B is

$$\frac{1}{2}\rho q_B^2 + \rho g h_B + p_B.$$
 (11.5–15)

Fig. 11-8 Since the liquid is open to the air at both A and B, we have approximately $p_A = p_B$. Moreover, approximately, $q_A = 0$. Hence equating the energy at A to that at B, we have finally

$$q_B = \sqrt{2g(h_A - h_B)}. (11.5-16)$$

The result embodied in eq. (11.5-16) is known as *Torricelli's theorem*. It is, however, only roughly approximate, for this reason: observation indicates that the shape of the stream flowing out of the hole is not a circular cylinder; rather it narrows down after leaving the orifice and becomes narrowest at C, a point called the "vena contracta." Between the opening and C the stream lines converge and it is only at C that they become approximately parallel and the pressure approximately equal to that of the atmosphere. The more accurate equation to replace (11.5-16) is then

$$q_C = \sqrt{2g(h_A - h_C)}.$$
 (11.5–17)

The size of the "vena contracta" depends on the type of opening. If the latter is merely a hole in a thin wall, the ratio of the area of the "vena contracta" to the area of the orifice is found by experiment to be approximately 0.62.

The discussion of the applications of Bernoulli's theorem has been so far for the case of an incompressible fluid. Application to a compressible fluid like a gas can, however, be handled if we revert to eq. (11·5–6) and evaluate the integral $\int \frac{dp}{\rho}$ by the substitution of the appropriate relation between p and ρ . If Boyle's law is satisfied we have

$$p = C\rho, \tag{11.5-18}$$

where C is a constant at any given constant temperature and therefore

$$\int \frac{dp}{\rho} = C \int \frac{d\rho}{\rho} = C \log \rho + C'$$

$$= C (\log p - \log C) + C', \qquad (11.5-19)$$

and eq. (11.5-6) then becomes

$$\frac{1}{2}\rho q^2 + \rho\Omega + \rho C \log p = C'', \qquad (11.5-20)$$

where C' and C'' are further constants. On the other hand if the gas law is the *adiabatic* one, which seems much more reasonable for gases moving with considerable velocities, we have, as the reader will recall from elementary physics,

$$p = K\rho^{\gamma}, \qquad (11.5-21)$$

where K is another constant and γ is the ratio of the specific heat of the gas at constant pressure to that at constant volume. Then

$$\int \frac{dp}{\rho} = K\gamma \int \rho^{\gamma-2} d\rho = \frac{K\gamma \rho^{\gamma-1}}{\gamma - 1} + K'. \quad (11.5-22)$$

The eq. (11.5-6) now becomes

$$\frac{1}{2}\rho q^2 + \rho\Omega + \frac{\gamma}{\gamma - 1}p = C''',$$
 (11.5–23)

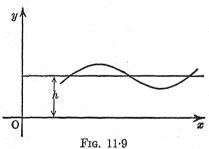
where C''' is a constant. It is seen that the same qualitative statement of Bernoulli's theorem holds for gases as for liquids. Hence the Pitot tube method may also be used to measure the velocity of a gas stream. Naturally the assumption of steady flow must still be made in spite of the fact that the presence of tubes, etc., introduces alterations in the flow. Hence the resulting equations become quasi-empirical and adjustment of constants necessary in the calibration of the appropriate apparatus.

The student should investigate for himself such qualitative

illustrations of Bernoulli's principle as the jet pump or aspirator and ship suction (the tendency of ships moving side by side to be drawn together), the Flettner rotor ship, the behavior of a ball in a jet, etc.

11.6. Waves in Fluids. In Chapter X we discussed wave propagation through a deformable medium and the general properties of wave motion. There are many important cases of waves in fluids, of which we shall consider only two, namely: (1) waves on the surface of a liquid, and (2) compressional waves through a liquid or gas. Indeed we shall still further specialize the first type as follows. Surface waves in a liquid are due to two causes, viz., the surface tension (see Sec. 11.8), and gravity. The former are of relatively short wave length, while the latter are very long. In our present treatment we shall confine our attention to the latter and particularize by considering the propagation of a gravity wave in a long straight canal.

Let the depth of the canal be h and suppose that it is small compared with the wave length. The displacement of the surface is assumed to take place in the xy plane and to be so small that, as in the case of the transversely vibrating string (Sec.



10·5), we may neglect the squares and higher powers of it and its derivatives. The x axis is taken as the horizontal bottom of the canal (see Fig. 11·9) so that the ordinate of the displaced surface is h + η, where η is the displacement in the y direction from the equilibrium

position. We shall assume that the pressure at any point (x, y) below the surface is the ordinary statical pressure. This means that we are supposing that the vertical accelerations of the particles of fluid are so small that they may be neglected. Hence for the pressure in question we have

$$p = \rho g(h + \eta - y), \qquad (11.6-1)$$

so that

$$\frac{\partial p}{\partial x} = \rho g \frac{\partial \eta}{\partial x} \cdot \tag{11-6-2}$$

Since from the equations of motion [eqs. $(11\cdot4-4)$] $\partial p/\partial x$ gives the horizontal acceleration, it follows that the latter is independent of y and hence is the same for all particles in the same vertical plane. Thus u is a function of x and t only. The equation of motion for the x direction (the external force being zero) is

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} = -\frac{1}{\rho} \frac{\partial p}{\partial x} \cdot \tag{11.6-3}$$

But since u is small, $u \frac{\partial u}{\partial x}$ is negligible compared with $\frac{\partial u}{\partial t}$, and (11.6-3) reduces to [using (11.6-2)],

$$\frac{\partial u}{\partial t} = -\frac{1}{\rho} \frac{\partial p}{\partial x} = -g \frac{\partial \eta}{\partial x}$$
 (11.6-4)

Now in this case $u = \frac{\partial \xi}{\partial t}$, where ξ is the *horizontal* displacement of the particles which were under equilibrium conditions in the plane at x. Hence we can write (11.6-4) in the form

$$\frac{\partial^2 \xi}{\partial t^2} = -g \frac{\partial \eta}{\partial x} \cdot \tag{11.6-5}$$

We must now apply the equation of continuity [eq. (11·3-8)], which in the present problem reduces to the two dimensional form

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = 0, (11.6-6)$$

so that

$$v = -\int_0^y \partial u \, dy = -y \, \frac{\partial u}{\partial x}, \qquad (11.6-7)$$

since $\frac{\partial u}{\partial x}$ is independent of y. Hence putting $v = \frac{\partial \eta}{\partial t}$ (at the sur-

face) and $u = \frac{\partial \xi}{\partial t}$, this becomes (if we recall that y = h at the surface)

$$\frac{\partial \eta}{\partial t} = -h \frac{\partial}{\partial t} \left(\frac{\partial \xi}{\partial x} \right) \cdot \tag{11-6-8}$$

Integrating with respect to the time yields

$$\eta = -h \frac{\partial \xi}{\partial x}, \qquad (11.6-9)$$

leaving off the constant of integration which is independent of t. If we utilize (11.6-9) in (11.6-5) we have finally the equation

$$\frac{\partial^2 \xi}{\partial t^2} = gh \frac{\partial^2 \xi}{\partial x^2}, \qquad (11 \cdot 6 - 10)$$

for the behavior of ξ . If we had eliminated ξ instead of η we should have found

$$\frac{\partial^2 \eta}{\partial t^2} = gh \frac{\partial^2 \eta}{\partial x^2} \cdot \tag{11.6-11}$$

We recognize at once that (11.6–10, 11) are examples of the one dimensional wave equation. Hence the disturbance on the surface is propagated with velocity

$$c = \sqrt{gh}. (11.6-12)$$

Suppose, for example, that the wave is of the simple harmonic variety. Then (Sec. 10.4)

$$\xi = \xi_0 \cos(\omega t - \kappa x), \qquad (11.6-13)$$

where $\omega = 2\pi\nu$, ν being the frequency of the wave and $\kappa = \omega/c$. The vertical displacement at the surface is then from (11.6-9)

$$\eta = -h\xi_0\kappa\sin(\omega t - \kappa x). \tag{11.6-14}$$

The elimination of the time between (11.6-13) and (11.6-14) gives the path of the liquid particle at the surface, viz.,

$$\frac{\xi^2}{\xi_0^2} + \frac{\eta^2}{h^2 \xi_0^2 \kappa^2} = 1. \tag{11.6-14a}$$

This is the equation of an ellipse with semi-major and semi-minor axes ξ_0 and $\xi_0 h \kappa$ respectively. Recalling that $\kappa = 2\pi/\lambda$, where λ is the wave length, and that $h \ll \lambda$, we see that the vertical axis is much shorter than the horizontal.

It is interesting to note that for h = 10 feet, the velocity of these surface waves is approximately 18 ft/sec.

The above analysis forms the basis for the investigation of the *dynamical* theory of the tides, which attributes the latter to wave motion in the water on the earth due to the attraction of the moon and sun. The details would carry us too far afield here.

We shall now turn to compressional waves in a fluid. These are of particular interest since all sounds are transmitted by waves

of this type (both in fluids and solids, as we have had occasion to notice). In this case the equations of motion become rather simple. Thus suppose within a fluid at rest a disturbance is created at some point. This disturbance will involve displacements of small elements of the fluid from their equilibrium positions. We shall assume that these displacements and the corresponding velocities are small enough so that their powers higher than the first and their products may be neglected. Hence the equations of motion may be written (for no external forces)

$$\frac{\partial u}{\partial t} = -\frac{1}{\rho} \frac{\partial p}{\partial x}, \quad \frac{\partial v}{\partial t} = -\frac{1}{\rho} \frac{\partial p}{\partial y}, \quad \frac{\partial w}{\partial t} = -\frac{1}{\rho} \frac{\partial p}{\partial z}.$$
 (11.6–15)

We shall now further assume the existence of a function $\phi(x, y, z, t)$ such that

$$u = -\frac{\partial \phi}{\partial x}$$
, $v = -\frac{\partial \phi}{\partial y}$, $w = -\frac{\partial \phi}{\partial z}$. (11.6–16)

The function ϕ is known as the *velocity potential*, and it is shown in more advanced treatises that its existence implies that the fluid has no angular velocity about any axis, i.e., its motion is *irrotational*. Substituting from (11.6-16) into (11.6-15) we have

$$\frac{\partial}{\partial t} \left(\frac{\partial \phi}{\partial x} \right) = \frac{1}{\rho} \frac{\partial p}{\partial x}, \quad \frac{\partial}{\partial t} \left(\frac{\partial \phi}{\partial y} \right) = \frac{1}{\rho} \frac{\partial p}{\partial y}, \quad \frac{\partial}{\partial t} \left(\frac{\partial \phi}{\partial z} \right) = \frac{1}{\rho} \frac{\partial p}{\partial z}. \quad (11.6-17)$$

Multiplying through these equations by dx, dy, dz respectively and adding, there results

$$\frac{\partial}{\partial t} d\phi = d \left(\frac{\partial \phi}{\partial t} \right) = \frac{1}{\rho} dp, \qquad (11.6-18)$$

where the differentials are, of course, space differentials. If we now integrate we obtain

$$\frac{\partial \phi}{\partial t} = \int \frac{dp}{\rho} + C, \qquad (11.6-19)$$

where we shall set C=0, since though it may be a function of t it cannot depend on x, y, z. Now the change in density is very slight in a region of the order of the maximum displacement. Therefore we can approximate by removing ρ from under the integral sign, replacing it by ρ_0 , the mean density of the fluid. The $\int dp$ then

becomes the variation of pressure at any point from the equilibrium pressure. We shall call it the excess pressure, and shall write it δp to avoid confusion with p which still denotes the absolute pressure. The eq. (11.6–19) then becomes

$$\frac{\partial \phi}{\partial t} = \frac{\delta p}{\rho_0} \cdot \tag{11.6-20}$$

There is always a relationship between the pressure and density of a fluid. We may for the moment ignore its exact nature and set merely

$$\delta p = c^2 \delta \rho, \qquad (11.6-21)$$

where $\delta \rho$ denotes the excess density in the fluid and c is at present a proportionality factor. We can then write (11.6–20) in the form

$$\frac{\partial \phi}{\partial t} = \frac{c^2}{\rho_0} \delta \rho. \tag{11-6-22}$$

There is considerable advantage to be obtained in our present study by introducing a new quantity called the *condensation*, defined as the ratio of the excess density to the mean density. Thus we have, denoting the condensation by s,

$$s = \frac{\delta \rho}{\rho_0} \cdot \tag{11.6-23}$$

One relation between condensation and velocity potential then is, from (11.6-22),

$$\frac{\partial \phi}{\partial t} = c^2 s. \tag{11.6-24}$$

Another relation between ϕ and s comes from the equation of continuity. If we write the latter in the form [eq. (11.3-6)]

$$\frac{\partial \rho}{\partial t} + \frac{\partial (\rho u)}{\partial x} + \frac{\partial (\rho v)}{\partial y} + \frac{\partial (\rho w)}{\partial z} = 0,$$

and substitute for ρ the expression $\rho_0(1+s)$, we have

$$\rho_0 \frac{\partial s}{\partial t} + \rho_0 (1+s) \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} \right) + \rho_0 \left(u \frac{\partial s}{\partial x} + v \frac{\partial s}{\partial y} + w \frac{\partial s}{\partial z} \right) = 0.$$
 (11.6–25)

Now experiments indicate that $\delta\rho$ is always very small compared with ρ_0 . Hence $s \ll 1$, and the variation of s in space is also very small. Therefore, from this and the approximations previously mentioned, the left-hand side of eq. (11-6-25) reduces to

$$\frac{\partial s}{\partial t} + \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0.$$
 (11.6–26)

If now we substitute from (11.6-16), the above becomes

$$\frac{\partial s}{\partial t} - \left(\frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} + \frac{\partial^2 \phi}{\partial z^2}\right) = 0. \tag{11.6-27}$$

This is the other relation between ϕ and s. We eliminate s between (11.6-24) and (11.6-27) and finally have

$$\frac{\partial^2 \phi}{\partial t^2} = c^2 \left(\frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} + \frac{\partial^2 \phi}{\partial z^2} \right) \cdot \tag{11.6-28}$$

It is shown in more advanced treatises that this is the differential equation of wave motion. The velocity potential is propagated throughout the fluid with velocity c. Indeed when ϕ is independent of y and z, (11.6-28) reduces to precisely the same form as (10.4-3) which, we saw, corresponds to the propagation of plane waves in a rod. It is interesting to note that the wave equation (11.6-28) here results from the combination of the equation of continuity and the equation of motion for small displacements from equilibrium, the first of which expresses the fact that a continuous, indestructible medium is involved and the second gives the slight motion from equilibrium of a small part of that medium conceived to move as a whole, viz., as a particle. There are other illustrations of this duality which emphasize its fundamental significance in wave motion.²

The waves of acoustics are all included in the above analysis. The two types of such waves which are most important in practice are respectively plane and spherical waves. In the former ϕ is

¹ See, for example, Jeans, "Electricity and Magnetism" (Cambridge University Press, 1925), p. 521. Also Love, "Mathematical Theory of Elasticity" (Cambridge University Press, 1920), Art. 210.

² See, for example, R. B. Lindsay, Proc. Nat. Acad. Sci., 17, 420, 1931.

a function of t and one space variable, say x, only, and eq. (11-6-28) reduces to the form

$$\frac{\partial^2 \phi}{\partial t^2} = c^2 \frac{\partial^2 \phi}{\partial x^2}.$$
 (11.6–29)

If we confine our attention here to plane harmonic waves as the most significant, we may write

$$\phi = A \cos(\omega t - \kappa x), \tag{11.6-30}$$

where ω and κ have the usual meaning and A is the velocity potential amplitude. The corresponding expressions for the particle velocity u, the particle displacement ξ , the excess pressure and the condensation may at once be obtained. Thus

$$\dot{\xi} = u = -\frac{\partial \phi}{\partial x} = -\kappa A \sin(\omega t - \kappa x),$$
 (11.6-31)

$$s = \frac{1}{c^2} \frac{\partial \phi}{\partial t} = -\frac{A\kappa}{c} \sin (\omega t - \kappa x), \qquad (11.6-32)$$

$$\xi = \frac{A}{c}\cos(\omega t - \kappa x), \qquad (11.6-33)$$

$$\delta p = \rho_0 c^2 s = -A \rho_0 c \kappa \sin \left(\omega t - \kappa x\right). \tag{11.6-34}$$

If we differentiate ξ with respect to x we obtain the important relation for acoustic wave transmission

$$s = -\frac{\partial \xi}{\partial x}.\tag{11.6-35}$$

Its physical meaning is as follows. If ξ increases in the positive x direction, s is negative corresponding to what may be called a rarefaction; similarly, decreasing ξ in the positive x direction corresponds to positive s and hence to a condensation. We ought to note that if ξ is always measured in the positive x direction then for a wave in the negative x direction, eq. (11.6-35) will become $s = +\frac{\partial \xi}{\partial x}$.

From eqs. (11.6-31, 32, 33, 34) we see at once that it is not only
$$\phi$$
 which satisfies the wave equation, but that u , s , and ξ are also solutions. Thus we have, for example

$$c^2 \frac{\partial^2 \xi}{\partial x^2} = \frac{\partial^2 \xi}{\partial t^2}$$
, and $c^2 \frac{\partial^2 s}{\partial x^2} = \frac{\partial^2 s}{\partial t^2}$. (11.6–36)

From the definition of phase (Sec. 10·4) we note that u, s, and δp are in the same phase, that is, when δp has a positive maximum, the same is true of u and s. However, while ξ and ϕ are in phase with each other, they are both out of phase with u, s, and δp .

Examining now the expression (11.6-21) we see that the general formula for the velocity of an acoustic wave in a fluid is

$$c = \sqrt{\frac{\delta p}{\delta \rho}} \cdot \tag{11.6-37}$$

The exact form of this will depend, of course, on the relation between p and ρ . If Boyle's law holds we have

$$c = \sqrt{\frac{p}{\rho}}.$$
 (11.6–38)

This formula (derived originally by Newton) does not, however, agree with the experimentally observed result, giving values which are too low. It seems more reasonable to assume that the relationship between p and ρ is adiabatic (as suggested by Laplace), viz., such that [eq. (11.5-21)]

$$\frac{\delta p}{\delta \rho} = \frac{\gamma p}{\rho}$$
.

We then have

$$c = \sqrt{\frac{\gamma p}{\rho}}, \qquad (11.6-39)$$

which agrees very well with the experimental values.1

All the general discussion of wave motion in Sec. 10.5 finds application in acoustic waves. But we may perhaps here refer the reader to some standard text on acoustics² where, for illustration, the interesting features of spherical waves in fluids are also discussed in detail.

11.7. Viscous Fluids. Our treatment of fluid motion has so far been confined to perfect fluids in which shearing strains never

¹ The student should check this by the substitution of numerical values. He may also show that in the case of a liquid the expression for the velocity is $c = \sqrt{k/\rho}$, where k is the bulk modulus.

² For example, P. M. Morse, "Vibration and Sound" (McGraw-Hill, New York, 2nd ed., 1948).

occur. It is hardly necessary to re-emphasize the *ideal* nature of the perfect fluid, and the fact that all fluids found in nature depart from this to a greater or lesser degree by displaying *viscosity*. Indeed in our short discussion of the kinetic theory of matter (Sec. 6.7) we noted that if a fluid is assumed to be composed of molecules in rapid motion there must be a tangential drag exerted on any layer of moving molecules by the contiguous layers. We there defined the coefficient of viscosity η as the tangential force per unit velocity gradient, and proceeded to find out a few interesting things about it in the case of a gas. For many problems of fluid flow like those treated in the previous sections the effect of viscosity is so slight that it may safely be neglected. However, there are some problems in which it is of considerable importance and we ought to mention one or two of these.

First let us observe that there is a connection between viscosity and the steady flow of a fluid. It will be recalled that in flow of this kind there exist stream lines which maintain their position as long as the flow lasts. Now in the flow of actual fluids it is found that while it is possible to have steady flow of a liquid through a tube for reasonably small velocities, in general as the velocity of flow increases there is a breakup of the stream lines and the motion may be said to become turbulent. There appears to exist for each fluid a certain critical velocity below which steady motion is possible but above which turbulence ensues. This may be readily shown, for example, by allowing a small amount of coloring matter to be transported in a flowing liquid. For small flow velocities it will be observed to follow a regular stream line, while eventually a velocity is attained for which the stream line is broken up and the coloring matter distributed generally throughout the tube. Now experiment indicates that for a viscous fluid this critical velocity is directly proportional to the coefficient of viscosity and inversely proportional to the density. Any theoretical investigation of the corresponding formula is beyond the scope of this book. We shall merely note that in any case this is a somewhat remarkable result, for it indicates that steady motion is possible for fluids of small viscosity only at very low flow velocities unless the density is also very small, while steady flow is possible over a considerable range of velocities for a very viscous fluid. We have an illustration of this in the steady flow of lava from a volcano.

Let us now discuss the steady flow of a viscous fluid through a

circular tube. We shall assume that the radius of the tube is R and that its axis lies along the x axis. We also suppose that the fluid is flowing with velocity u in the positive x direction. Consider a coaxial cylindrical shell whose inner surface has the radius r while the outer surface has the radius r+dr. Due to the viscosity there will be a tangential stress per unit area of the inner surface of magnitude

$$X_r = \eta \frac{du}{dr},\tag{11.7-1}$$

where η is the coefficient of viscosity and $\frac{du}{dr}$ is the velocity gradient normal to the axis of the tube. This follows at once from the definition of η . Then the tangential force on the whole inner surface for a length l of tube is

$$2\pi r l \eta \, \frac{du}{dr} \, \cdot \tag{11.7-2}$$

The corresponding tangential force on the outer surface of the shell will then be

$$2\pi r l \eta \frac{du}{dr} + \frac{d}{dr} \left(2\pi r l \eta \frac{du}{dr} \right) dr. \tag{11.7-3}$$

There will thus be a net tangential force on the shell in the positive x direction of

$$\frac{d}{dr}\bigg(2\pi rl\eta\,\frac{du}{dr}\bigg)\,dr.$$

The flow is supposed to be steady. Hence the above force must be equilibrated by a force due to the difference in pressure $p_l - p_0$ at the two ends of the tube. The force due to the latter will be $(p_l - p_0)2\pi r dr$; equating the two equal forces gives the differential equation

$$\frac{d}{dr}\left(2\pi r \ln \frac{du}{dr}\right) = (p_l - p_0)2\pi r. \tag{11.7-4}$$

A first integration yields

$$2\pi r l \eta \frac{du}{dr} = (p_l - p_0)\pi r^2 + C_1, \qquad (11.7-5)$$

while the second gives

$$u = \frac{(p_l - p_0)r^2}{4\eta l} + \frac{C_1 \log r}{2\pi\eta l} + C_2.$$
 (11.7-6)

We must now suppose that at the boundary of the tube the viscosity causes the fluid to be at rest. Hence u=0 for r=R. Moreover, the value of u for r=0, i.e., along the axis, must be finite. Hence the coefficient of $\log r$ must vanish. This means $C_1=0$. From the previous condition we get

$$C_2 = -\frac{(p_l - p_0)R^2}{4nl},$$

and we finally have for the velocity u at any distance r from the axis

$$u = \frac{(p_l - p_0)(r^2 - R^2)}{4\eta l}.$$
 (11.7-7)

We may at once calculate the rate of volume flow per second passing any cross-section of the tube by integrating $2\pi ru dr$ from r=0 to r=R. We get

$$\int_0^R 2\pi r u \, dr = \frac{\pi R^4}{8\eta} \cdot \frac{(p_0 - p_l)}{l} \,. \tag{11.7-8}$$

This is known as *Poiseuille's* formula. In words it states that the volume rate of steady flow of a viscous fluid through a cylindrical tube varies directly as the fourth power of the radius and directly as the pressure gradient, while inversely as the coefficient of viscosity. For very small tubes the velocity may thus become very small. Poiseuille verified the formula by experiments on the flow through capillary tubes. Incidentally the formula may be used to determine the coefficient of viscosity.

As might be expected, the passage of an acoustic wave through a viscous fluid is accompanied by an attenuation of the intensity of the wave. This is particularly marked in the case where the transmission takes place through a very narrow tube.

Some of the most important problems in applied physics are connected with the resistance experienced by a solid object in moving through a viscous fluid. We shall mention but one illus-

¹ See, for example, Stewart and Lindsay, "Acoustics," p. 67 ff.

tration. By means of analysis beyond the scope of this book Stokes showed that the resisting force on a sphere of radius a moving through a medium of viscosity η with constant velocity v is given by

$$F = 6\pi \eta a v, \qquad (11.7-9)$$

which the reader may be willing to accept as physically plausible, at any rate. This formula is fundamental for the work of Millikan on the measurement of the charge on the electron by the behavior in electrostatic fields of very small electrically charged drops of oil. Thus consider an oil drop of radius a (actually in the neighborhood of 10^{-4} cm) and of density ρ_0 falling under the action of gravity in a fluid of density ρ at rest. There are then three forces acting on the drop, viz., gravity, of magnitude

$$\frac{4}{3}\pi a^3 \rho_0 q$$
:

the buoyancy, of magnitude (by Archimedes' principle)

$$\frac{4}{3}\pi a^3 \rho g$$
;

and the *resistance* of the fluid given by the law of Stokes above. In the steady state when the velocity is constant, the balance of forces requires that

$$\frac{4}{3}\pi a^3(\rho_0 - \rho)g = 6\pi \eta a v, \qquad (11.7-10)$$

so that the velocity of fall is therefore

$$v = \frac{2}{9}a^2 \frac{(\rho_0 - \rho)}{\eta} g. \tag{11.7-11}$$

In the experiment to determine the charge on the electron, a charge is given to the drop by illuminating the region in which it moves with X-rays or ultraviolet light. Suppose that this charge is e, the smallest possible charge, i.e., that of the electron. An electric field is then applied to the space where the drop is moving and in the vertical direction. If the intensity of the field is E, i.e., E dynes per unit charge, the total force on the charged drop due to the field is eE and if E is adjusted until the drop comes to rest we must have for equilibrium

$$eE = \frac{4}{3}\pi a^3(\rho_0 - \rho)g.$$
 (11.7-12)

If we could measure all quantities here we could then solve for e. However, a is very difficult if not impossible to measure accurately.

Fortunately we can solve for it from eq. (11.7-11) and then substitute into eq. (11.7-12), finally obtaining

$$e = \frac{18\pi}{\sqrt{2}E} \frac{(v\eta)^{3/2}}{\sqrt{g(\rho_0 - \rho)}}.$$
 (11.7-13)

As a matter of fact the values of e obtained by this formula, though of value as first approximations, are found to depend on the pressure of the air and hence can not be exact. The drops are so small that the molecular constitution of the air must be taken into consideration. That is, Stokes' law, which is based on the assumption that the fluid in question is perfectly continuous, must be modified. We shall not enter into further discussion here.

11.8. Surface Phenomena. Capillarity. It would hardly be appropriate to leave the subject of fluids without some elementary reference to the phenomena associated with liquid-gas and liquid-solid interfaces. These phenomena are of great importance in applied physics and physical chemistry.

It is a common observation that the surface of a liquid exposed to the atmosphere appears to act like a more or less tight skin or membrane covering the body of the liquid. This property is manifested in a variety of ways. It is very evident, for example, in the apparent violation of the laws of hydrostatics shown in the floating of a needle or particles of sand on the surface of water. It is also displayed in the characteristic form of liquid drops, all tending toward the spherical shape, the sphere having the smallest surface enclosing a given volume. It is as if the surface tries to contract as much as possible. The student will recall other illustrations of the same nature from elementary physics.

Another common observation is the rise of certain liquids in tubes of small bore, called *capillary* (i.e., hair-like) tubes. Indeed it is this name which has become associated with the whole subject under discussion.

In beginning our survey of these phenomena we shall assume that if a straight line is drawn in the surface of a liquid the portion of the surface on one side of the line exerts on that on the other side a force which is directly proportional to the length of the line.

¹ See R. A. Millikan, "Electrons (+ and -), Protons, Photons, Neutrons, Mesotrons, and Cosmic Rays" (Chicago, Rev. Ed., 1947), p. 90 ff.

The coefficient of proportionality or force per unit length is called the surface tension, which we shall denote by the symbol T_s . It is dependent on the nature of the liquid and the substance above its surface, and also on the state of the surface and the temperature, decreasing with increase of the latter. It is, however, independent of the size of the surface, differing therein from the superficial tension in an elastic sheet or membrane. For the latter, being due to the action of external forces, increases with the further stretching of the surface.

Perhaps the clearest way to visualize the meaning of surface tension is to consider a simple experiment which, if done carefully enough, is competent to give an approximate value of T_s for the

case of several liquids. A piece of perfectly clean thin metal wire (e.g., platinum) is bent into the form of a rectangular framework (Fig. 11·10) and suspended in the liquid from one arm of a balance. As the metal frame is pulled out of the liquid with a two faced film attached to the sides of the frame it is necessary to balance the tendency of this film

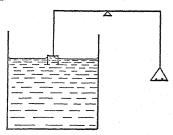


Fig. 11.10

to contract by placing weights in the balance pan. When equilibrium is attained it will be found that the counterpoising weight is independent of the magnitude of the surface film formed. The value of the surface tension may then be computed from the simple equation

$$2l T_s = mq, (11.8-1)$$

where mg is the counterpoising weight and l is the horizontal length of the film. The factor 2 enters from the fact that the film has two surfaces. Of course it must be noted that the extension of film surface possible without breaking varies a good deal with the liquid used, and the values obtained in this way vary greatly with the cleanness of the surface. The slightest contamination, as for example with oil from hair or fingers, suffices to reduce the value of T_s materially.

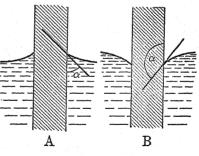
An understanding of the nature of surface tension may be gained if we examine the phenomenon in terms of the molecular theory which considers the liquid to be composed of a large number of

molecules (much like those of a gas — Secs. 6.6 and 6.7) moving about with considerable velocity, but also exerting on each other attractive forces which are rather small when the molecules are at ordinary distances, but which become extremely large when the molecules are very close to each other. Certain evidence makes it probable that this sphere of molecular action, as it is called, is usually of the order of 10⁻⁷ cm, hence about ten times the order of magnitude of molecular dimensions. In the interior of a liquid a given molecule will be surrounded by other molecules on all sides and hence be under the influence of a force which is approximately the same in every direction. This, however, is not the case with a molecule in the surface of the liquid, for the attraction of the gas molecules above the surface will be far outweighed by the attraction of the molecules of the liquid. Hence there will act on such a molecule a resultant downward force, and it is this force which we may look upon as producing the effect of a tight skin on the surface. That this is indeed only an approximate picture of the phenomenon will be recognized from the observed fact that on the molecular theory it is necessary to assume that conditions on the surface of a liquid in contact with a gas containing vapor of the liquid are by no means static, since rapidly moving molecules of the liquid are continually escaping into the region above the liquid, and concomitantly vapor molecules are flying back into the liquid. reader will recall that when the average number traveling in each direction per unit of time is the same the vapor above the liquid is said to be saturated. If the number escaping is greater than the number regained there is a net evaporation of the liquid into vapor form, while if the tendency of the process is in the other direction we speak of the condensation of the vapor. By reason of this state of affairs it is perhaps better to consider the surface phenomena from the standpoint of energy. The kinetic energy of a molecule moving entirely in the interior of the liquid is affected only by collisions with its neighbors, but when such a molecule approaches the surface it must lose kinetic energy because of the effective force tending to pull it back. It will then gain potential energy, so that the molecules in the surface will have on the average greater potential energy than those in the interior. Now we have already had occasion to notice that the most stable equilibrium of an aggregate of particles corresponds to the least potential energy compatible with the constraints, etc., (Sec. 5.8). Hence the surface will so behave as to keep the total potential energy as small as possible. This may be brought about, however, by keeping the area of the surface as small as possible. Hence the tendency which gives the effect of a tight membrane over the surface. This moreover accounts for the previously mentioned tendency for drops to assume the spherical form. The foregoing molecular discussion also accounts qualitatively for the observed fact that the surface tension decreases with increase of the temperature.

The considerations above may also be applied in an interesting manner to the stretching of the surface of a liquid. This involves the transfer of molecules from the interior to the surface with consequent increase in potential energy and decrease in the kinetic energy. But, as we have seen, the temperature depends on the average value of the latter. Hence if the temperature is not to fall, heat must be absorbed by the surface during the stretching. This is in addition to the mechanical work done against the surface tension during the process of expansion. The mechanical equivalent of the heat absorbed plus the mechanical work done per unit increase in area will be called the total surface energy of the liquid. The reader will find no difficulty in showing that the mechanical work per unit increase in area is numerically equal to the surface

tension. At any temperature (at which the liquid exists as such) the surface energy is always numerically greater than the surface tension.

When a solid object is brought into contact with a liquid so as to be partially immersed, it is observed that one of two things will happen: either the liquid will appear to run a certain distance up



Frg. 11-11

the side of the object as in A (Fig. 11·11) or will appear to be depressed somewhat in the vicinity of the object as in B. In the first case the liquid is said to wet the object, while in the second case it does not wet the object. To put the matter somewhat crudely we may say that there is an attraction between the liquid molecules and those of the object, which in the immediate neighborhood of the object (in case the latter is wet) outbalances the

mutual attraction of the liquid molecules. This, for example, is true when the liquid is pure water and the solid is glass. On the other hand, in mercury which does not "wet" glass, the mutual action of the liquid molecules is greater than the attraction of the mercury-glass molecules and hence the mercury surface is depressed in the neighborhood of a glass plate. The angle between the surface of the liquid and the surface of the solid is known as the angle of contact — represented by α in Fig. 11-11. It varies with the cleanness of the surfaces, but experiment indicates that it is fairly definite for a given pair of clean surfaces. For example in the case of mercury in contact with glass it is about 140°, while for water in contact with glass it is practically zero.

These considerations are very approximate, and we must emphasize that capillary phenomena have been the subject of much

Fig. 11·12

theorizing extending from the time of Laplace and (later) Gauss down to the present period. The subject is still incomplete owing to our ignorance of the precise nature of the cohesive forces between molecules both of the same and different kinds.

Nevertheless we can derive a few more elementary results. Consider first the rise of a liquid in a capillary tube as indicated in Fig. 11·12. Let the radius of the tube be a and suppose it is placed in a liquid of density ρ which wets it. Let the angle of contact be α (an acute angle) and assume that the surface tension of the liquid in contact with air is T_s . Where the

liquid is in contact with the tube there will then be a force whose vertical component acts upward and has the magnitude

 $2\pi a T_s \cos \alpha$.

The liquid will therefore rise in the tube until this force is balanced by the weight of the column of liquid above the surface of the liquid into which the tube has been placed. Since the upper surface of the liquid in the tube is not level, the height of the column is not a definite quantity, but we can nevertheless take for it the average height (i.e., that which it would have if it were level) and call this h. The weight of the column then becomes $\rho g h \pi a^2$.

and hence equating this to the former quantity, we get

$$2\pi a T_s \cos \alpha = \rho g h \pi a^2, \qquad (11.8-2)$$

or

$$h = \frac{2T_s \cos \alpha}{\rho g a}, \qquad (11-8-3)$$

for the height, which thus proves to be inversely proportional to the radius of the tube. If $\cos \alpha$ is negative there is no capillary rise, but a depression instead. This occurs in the case of mercury in glass, where $\alpha=140^\circ$, as has already been mentioned. The law given by eq. (11.8–3) is known as Jurin's law. It can be used to determine T_s by an alternative method to that described in the first part of the section. The reader may show that the same

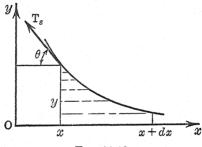


Fig. 11.13

law applies to the capillary rise between two vertical plates placed in a liquid which wets them, provided that a now refers to half the distance between the plates.

The form of the liquid surface at rest under gravity in the neighborhood of a solid object can under certain conditions be determined very readily by an elementary method due to Maxwell. This is the case, for example, when the object is a vertical flat plate of great breadth. We assume that the object is wet by the liquid. Let us take a section of the surface (considered as a cylinder) in the xy plane (Fig. 11·13). It is understood that for y = 0 the surface is plane and the section parallel to the x axis. Consider the infinitesimal slice of unit thickness whose projection

on the xy plane is the area between the ordinates corresponding to x and x + dx. Since this portion is in equilibrium the total force on it in the x direction, say, must vanish. First let us consider the force due to the surface tension. At the left the horizontal component is (since T_s is the force per unit length)

$$-T_s \cos \theta = -\frac{T_s}{\sqrt{1 + \left(\frac{dy}{dx}\right)^2}}, \qquad (11.8-4)$$

where dy/dx is the slope of the curve at the point in question. At the right, on the other hand, the horizontal component is

$$T_s \left[\cos \theta + \frac{d \cos \theta}{dx} \, dx \right]$$
 (11.8-5)

The net horizontal component is the sum of (11.8-4) and (11.8-5) or

$$T_s \frac{d \cos \theta}{dx} dx,$$

which from (11.8-4) is equal to

$$\frac{-T_{s}\frac{dy}{dx}\frac{d^{2}y}{dx^{2}}dx}{\left[1+\left(\frac{dy}{dx}\right)^{2}\right]^{3/2}}.$$
(11-8-6)

This force must be balanced by the hydrostatic forces against the two vertical ends of the element. If we assume for simplicity that there is a vacuum above the surface, then the pressure at any point in the portion of the liquid which we are considering will be zero or negative. Thus for any value of y, it will be

$$-\rho gy$$
.

Hence the total thrust on unit breadth of the end of the element at x is

$$-\rho g \int_0^y y \, dy = -\frac{\rho g}{2} y^2. \tag{11.8-7}$$

At x + dx, the corresponding thrust in the *opposite* direction is

$$-\frac{\rho g}{2}\left(y + \frac{dy}{dx}\,dx\right)^{2} = -\frac{\rho g}{2}\left(y^{2} + 2y\,\frac{dy}{dx}\,dx\right), \quad (11.8-8)$$

if we neglect differentials of order higher than the first. Hence the net thrust due to hydrostatic pressure is

$$-\rho gy \frac{dy}{dx} dx. \tag{11.8-9}$$

Since the surface tension force and the hydrostatic thrust balance, we must have

$$\frac{T_s \frac{d^2 y}{dx^2}}{\left[1 + \left(\frac{dy}{dx}\right)^2\right]^{3/2}} = \rho gy.$$
 (11-8-10)

From analytic geometry we recall that

$$\frac{\frac{d^2y}{dx^2}}{\left[1+\left(\frac{dy}{dx}\right)^2\right]^{3/2}}=\frac{1}{R},$$



where R is the radius of curvature of the surface at the point in question. Hence eq. (11.8-10) becomes

$$\rho gy = \frac{T_s}{R} \cdot \tag{11.8-11}$$

Strictly speaking there should be a constant of integration added to ρgy to give the value of T_s/R for y=0. But since we have assumed that the surface becomes a plane when y=0 (i.e., at considerable distance from the immersed plate), we have $R \doteq \infty$ for y=0, and hence the constant will vanish. Therefore (11.8–11) is correct in this case as it stands. It describes the form of the surface in terms of the radius of curvature. It is, however, possible to put it into more usable form if we multiply through (11.8–10) by $\frac{dy}{dx} dx$ obtaining

$$\rho gy \ dy = -T_s d \left[\frac{1}{\sqrt{1 + \left(\frac{dy}{dx}\right)^2}} \right], \qquad (11.8-12)$$

as may be verified immediately. Integration then yields

$$\frac{1}{2}\rho g y^2 = -\frac{T_s}{\sqrt{1 + \left(\frac{dy}{dx}\right)^2}} + C.$$
 (11.8–13)

Now $dy/dx = \tan \phi$, where ϕ is the slope of the curve (i.e., the supplement of θ), and hence we may write the above [recall eq. (11.8-4)]

$$\frac{1}{2}\rho g y^2 = T_s \cos \phi + C. \tag{11.8-14}$$

Now when $y=0,\,\phi=\pi$ and hence $C=T_s,$ so that

$$\frac{1}{2}\rho g y^2 = T_s (1 + \cos \phi), \qquad (11.8-15)$$

or if we prefer to use θ ,

$$\frac{1}{2}\rho g y^2 = T_s (1 - \cos \theta), \qquad (11.8-16)$$

which is a somewhat more convenient form than (11.8–11). The complete equation in cartesian notation is rather complicated and will not be given here. We may note, however, that the same curve is encountered in the stretching of a uniform spring. It may also be pointed out that the form of the surface is the same as that of the top of a liquid drop resting on a flat horizontal plate which it does not wet. It should be emphasized that the general problem of finding the capillary surface can not be handled by elementary methods and is extremely difficult.

The weight of the liquid raised by the capillary action is naturally equal to the volume above the plane level multiplied by ρg . Another expression for the same thing is, however, to be found by evaluating the total *vertical* component of the forces acting on the element considered in our previous discussion. This is

$$T_{s} \frac{d \sin \theta}{dx} dx = T_{s} \cos \theta \cdot d\theta. \qquad (11.8-17)$$

Hence the total weight is simply

$$T_s \int_0^{\theta} \cos \theta \, d\theta = T_s \sin \theta. \tag{11.8-18}$$

¹ Thomson and Tait, "Treatise on Natural Philosophy" (Oxford University Press, 1867), Vol. 1, p. 455.

When two parallel vertical flat plates are placed in a liquid the capillary rise or depression produces a force tending to cause

the plates to attract or repel each other. We can evaluate the magnitude of this force in any case by the simple analysis just given. Thus consider the case shown in Fig. 11·14 where the plates are separated by a distance l and the angles of contact are respectively α_1 and α_2 . On the left the liquid is drawn up to a height h_1 above the low level outside the plates and at the right to the height h_2 . Considering first the left plate,

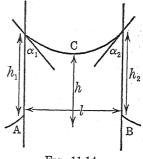


Fig. 11-14

we see that there is a force per unit breadth of plate acting at A tending to pull it towards the right of magnitude

$$T_s \sin \alpha_1 + \frac{1}{2} \rho g h_1^2,$$

as is clear from the immediately preceding analysis. On the other hand acting at C there is a force T_s acting to pull it towards the left. The net force toward the right is therefore

$$T_s(\sin \alpha_1 - 1) + \frac{1}{2}\rho g h_1^2.$$
 (11.8–19)

Now to go back to eq. (11.8-14), it develops that since y = h (see Fig. 11.14) for $\phi = \pi$, we have $C = T_s + \frac{1}{2}\rho gh^2$ and hence

$$\frac{1}{2}\rho g h_1^2 = \frac{1}{2}\rho g h^2 + T_s (1 + \cos\phi_1)
= \frac{1}{2}\rho g h^2 + T_s (1 - \sin\alpha_1).$$
(11.8-20)

Hence the net force toward the right on the left plate becomes at once

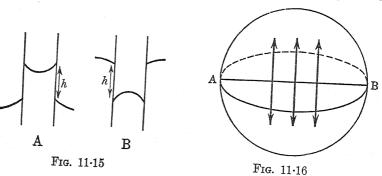
$$\frac{1}{2}\rho gh^2$$
, (11.8–21)

and a precisely similar expression results for the net force on the right-hand plate toward the left. Here h is the height of the bottom of the meniscus above the level of the liquid outside the two plates. In the case where the plates are both of the same material and close together $\alpha_1 = \alpha_2 = \alpha$, and we can apply eq. (11.8-3) for the capillary tube. Then the force pulling the plates

together is approximately

$$\frac{1}{2}\rho g \cdot \frac{4T_s^2 \cos^2 \alpha}{\rho^2 g^2 a^2} = \frac{2T_s^2 \cos^2 \alpha}{\rho g a^2}, \qquad (11.8-22)$$

where here a is half the distance between the plates. The force of attraction between two nearby plates immersed in a liquid which wets both of them is therefore inversely proportional to the square of the distance between them. This is to be sure an approximate result. It will not apply, for example, when one of the plates is



wet and the other not; in the latter case a repulsion rather than an attraction is observed to take place. 1

From our study of the capillary rise it is clear that the pressure at a point immediately underneath the curved capillary surface in a tube is less or greater than the pressure in the gas above the surface according as the liquid rises or is depressed in the tube (Fig. 11·15, A, B). The amount of pressure difference is given at once by ρgh , where h is the rise or depression respectively. Hence employing the expression for h given in eq. (11·8–3) we have for the pressure difference

$$\Delta p = \frac{2T_s \cos \alpha}{a}, \qquad (11.8-23)$$

i.e., inversely proportional to the radius of the tube. This result suggests at once a more general one. Consider a sphere of liquid of radius R (Fig. 11·16). Draw any diametral plane AB. This will cut the spherical surface in a great circle of radius R, and about

¹ The calculation in the latter case can be shown to follow from the usual analysis. See Prob. 21 at the end of the chapter.

the periphery of this circle there will be a tension of T_s dynes per cm or a total force of $2\pi RT_s$ dynes tending to pull the two hemispheres (formed by the diametral plane) together. This must be balanced by an excess pressure (over the atmospheric) inside the sphere, which will exert a force on the diametral plane of magnitude $\Delta p \cdot \pi R^2$. Equating the two balanced forces we have

$$\Delta p = \frac{2T_s}{R} \cdot \tag{11.8-24}$$

If instead of a sphere of liquid we have a hollow spherical film, such as a soap bubble, a similar result follows, except that since there are *two* surfaces to the film we now have

$$\Delta p = \frac{4T_s}{R} \tag{11.8-25}$$

for the excess pressure inside the bubble. Of course, in this case the student will need to prove that the pressure, which acts everywhere normally to the inner surface of the film produces an effective force $\Delta p \cdot \pi R^2$ across any diametral section (see Prob. 20 at the end of the chapter).

More elaborate analysis than we shall attempt here indicates that the pressure at any given point just inside a liquid surface due to the surface tension (i.e., the excess or defect over the pressure outside) is given by the formula

$$\Delta p = T_s \left(\frac{1}{R_1} + \frac{1}{R_2}\right),$$
 (11.8-26)

where R_1 and R_2 are the so-called *principal radii* of curvature of the surface at the point considered. It is at once evident that for the special case of the sphere where $R_1 = R_2 = R$, the radius of the sphere, for all points, eq. (11.8-26) reduces to the previously derived simple form (11.8-24).

The accurate determination of the surface tension for different liquids under different physical conditions is a very important matter for modern physical chemistry.¹ We shall stop to notice

¹ See, for example, Willows and Hatschek, "Surface Tension and Surface Energy," (2nd edition, 1919, Philadelphia). A more recent and very complete study of surface phenomena will be found in H. Freundlich, "Colloid and Capillary Chemistry" (Dutton, New York, 1926). See also E. K. Rideal, "Introduction to Surface Chemistry" (Cambridge University Press, 1930).

but one of the many interesting modern problems connected with surface effects. Study of the behavior of solutions of chemical salts has shown that in most of these there is a difference between the concentration (i.e., the mass of dissolved salt per unit volume of the solvent) in the surface layer and in the interior portion of the solution. This phenomenon is known as adsorption, and we may speak of positive adsorption if the concentration is greater in the surface than in the bulk, or negative adsorption if the reverse. is true. Now it is also found that, as might be expected, a dissolved substance alters the surface tension of the solvent. Thus for example it is observed that in the case of NaCl the solution of each gram-equivalent mass of salt in water increases the surface tension by 1.53 dynes/cm. On the other hand some substances lower the surface tension on solution. There is a close connection between the effect of concentration on the surface tension and the adsorption. Thus if u denotes the excess of dissolved substance (or solute) in the surface in grams/cm2 over the bulk concentration, called C, the following formula first derived by Willard Gibbs,

$$u = -\frac{C}{RT} \cdot \frac{dT_s}{dC} \tag{11.8-27}$$

holds, where T is the absolute temperature and R is the gas constant appropriate to the solution. Since many chemical reactions are found to take place most readily at the surfaces of substances it is clear that adsorption plays an extremely important rôle in chemistry.

For further information on modern aspects of surface tension phenomena, reference may be made to the comprehensive treatment of Freundlich in the text mentioned in the footnote. An interesting experimental study of the properties of soap films may be found in *Soap Films*, by A. S. C. Lawrence (G. Bell and Sons, London, 1929).

PROBLEMS

1. A vessel in the form of a rectangular parallelepiped of dimensions a, b, and c is filled with water to a height h. Calculate the total thrust due to the water on the bottom and each side of the vessel. Find the center of pressure on each side, i.e., the point where the resultant thrust on the side acts. Hint:

 $^1\,\mathrm{For}$ the derivation of (11·8–27) the student may consult Willows and Hatschek, loc. cit., p. 109 f.

If the x axis is taken as the line of intersection of the surface with one side and the y axis is the line perpendicular to this (i.e., parallel to the depth), show first that the coördinates of the center of pressure will be given by

$$x_c = \frac{\int \int x p \ dx \ dy}{\int \int p \ dx \ dy},$$

and

$$y_c = \frac{\int \int yp \ dx \ dy}{\int \int p \ dx \ dy},$$

where p = f(x,y) as indicated in the text. The integration is to be carried out over the whole area in contact with the liquid.

- 2. Calculate the total thrust on an equilateral triangle of side a immersed vertically in a liquid of density ρ with the base parallel to the surface of the liquid and at distance h from the surface. Also compute the position of the center of pressure.
- 3. A hollow sphere of radius R is completely filled with water. What is the resultant vertical thrust on the inside surface? Compare the vertical thrusts on the upper and lower halves of the surface.
- 4. A cylindrical vessel of length l and radius a with plane ends is filled with water and rests with its axis horizontal. Compare the vertical thrusts on the upper and lower halves of the curved surface. Also calculate the total thrust on each end and determine the position of the center of pressure at each end.
- 5. A sphere floats with $\frac{1}{3}$ of its surface above water. Find its mean density.
- 6. A cylindrical can 30 cm long and 10 cm in diameter and weighing 100 grams cannot float stably in water in an upright position. Find the amount of mercury (density 13.6 grams/cm²) which must be placed in the can in order that it may just be able to float upright stably. How far will the can sink in the water?
- 7. Assuming Boyle's law, derive the law of variation of atmospheric pressure with height above the earth's surface.
- 8. Assuming the adiabatic law, derive the law of variation of atmospheric pressure with height above the earth's surface.
- 9. If the density in a certain liquid at rest varies linearly with the depth, find the expression for the pressure as a function of depth.
- 10. Find the total force that the air exerts on the walls, ceiling, and floor of a room 10 meters by 8 meters by 3 meters.
- 11. Prove that if an incompressible homogeneous fluid moves so that its velocity components satisfy the relations u = kx/r, v = ky/r and w = kz/r, where $r = \sqrt{x^2 + y^2 + z^2}$ and k is a constant with the dimensions of velocity, the equation of continuity is satisfied. Interpret this type of motion physically.

- 12. Derive Bernoulli's theorem (eq. 11.5-7) by assuming that the total energy per unit volume along any stream line is constant (i.e., conservation of energy applied to fluid motion). Hint: Recall that the potential energy of a fluid per unit volume due to pressure is equal to the pressure (Sec. 11.4).
- 13. Water flows out of an orifice at the bottom of a vertical tank (see Fig. 11-8). If the area of cross section of the orifice is S_B and that of the tank is S_A , use Bernoulli's theorem and the equation of continuity to calculate the theoretical rate of flow out of the orifice and the rate of descent of the free surface in the tank. (Introduce numerical values to get order of magnitude.)
- 14. Use the result of the preceding problem to obtain the height of the free surface in the tank above the orifice at any time t after the flow has begun and draw a graphical sketch of the dependence of descent on the time.
- 15. How many gallons of water flow in 24 hours through a pipe 4 in. in diameter if at a constriction of 1.5 in. in diameter a mercury pressure gauge shows a difference in height of 1.8 in. from another in the unconstricted portion of the pipe?
- 16. Show how the Pitot tube may be used to measure the velocity of a rapid blast of air in a tube (such as might be produced by a blower). (Eq. 11.5-23.)
- 17. Deduce the dependence of the velocity of sound in a gas on the temperature and pressure.
- 18. By setting up the boundary conditions (continuity of pressure and continuity of volume displacement) at the interface between two fluid media, compute the relations between incident and reflected amplitudes and incident and transmitted amplitudes of an acoustic wave incident normally on the interface. (Take harmonic waves.) Discuss *phase* relations.
- 19. Defining the intensity of an acoustic wave as in Sec. 10·4, show that the intensity of a plane acoustic wave in a fluid may be written in the form

$$I=\frac{1}{2}\frac{p^2_{\max}}{\rho_0 c},$$

where p_{max} is the maximum excess pressure in the fluid and ρ_0 and c have their usual significance. Comment on the connection with eq. 10-4-45.

- 20. Prove that the excess pressure inside a soap bubble produces an effective force across any diametral plane equal to the pressure times the area of a great circle of the sphere. (Sec. 11.8, eq. 11.8–24, etc.)
- 21. Two plane plates are introduced close together into a liquid, which wets the one but not the other. Show that the two plates will repel each other and calculate the magnitude of the repulsion, following the analysis of Sec. 11.8.
- 22. Prove that Jurin's law (11.8-3) applies equally well to the capillary rise between two parallel plates placed in a liquid which wets them.
- 23. Prove that when n equal small spheres of water coalesce so as to form a single drop the surface energy is diminished in the ratio $n^{-1/3}$.

24. What form does the one-dimensional equation of motion of a perfect fluid confined by rigid walls, viz.,

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} = -\frac{1}{\rho} \frac{\partial p}{\partial x},$$

take if the fluid is incompressible? If the pressure gradient is due entirely to gravity, find the acceleration of the fluid at any point and compare with Torricelli's theorem.

25. The wave length of a long gravity wave on water of depth 20 feet is 100 feet. If the wave amplitude parallel to the water surface is 2 feet, find the maximum value of the total particle displacement. Do the same for the total particle velocity and compare with the wave velocity.

26. A plane harmonic acoustic wave in the x direction has the particle displacement in the form $\xi=A$ cos $(\omega t-kx)$, where $k=\omega/c=2\pi\nu/c$ and ν is the frequency. Find the expressions for the velocity potential ϕ , the condensation and excess pressure.

27. In Problem 26, if $\nu=512$ cycles/sec and the wave is in air at 20° C, find A if the excess pressure amplitude is 10^{-1} dynes/cm². Also find the condensation amplitude.

28. Prove that the potential energy density in an acoustic wave in a fluid is $\frac{1}{2} \rho_0 c^2 s^2$. Show that for a plane wave the average potential energy density equals the average kinetic energy density. Show that the intensity of a plane wave is given by the product of the average energy density and the wave velocity.

29. In Problem 18 of Chapter X the damping coefficient R in the wave equation for acoustical waves in a fluid becomes $4\omega^2\eta/3\rho_0$, if the damping is due to viscosity (coefficient of viscosity = η and the other symbols have their usual meaning). How far will a 100-cycle wave travel in air before its amplitude is decreased in the ratio 1/e? Work the same problem for water. Take $\eta_{\rm air} = 1.8 \times 10^{-4}$ dyne sec/cm² and $\eta_{\rm water} = 1.14 \times 10^{-2}$ dyne sec/cm².

CHAPTER XII

ADVANCED MECHANICS

12.1. Hamilton's Principle. In the previous chapters of this book we have based the various applications of mechanics for the most part on the principles set forth in Chapter I, namely, the Newtonian "laws" of motion. We did indeed emphasize in Chapter VIII that there exist alternative wavs of expressing the fundamental principles and we gave as examples D'Alembert's principle and the Gaussian principle of least constraint. It is important to realize that certain problems may be handled in a more direct and efficient fashion by some one principle than by others. progress of mechanics has been marked by a continual search for new points of view. We wish in this chapter to discuss one which stresses primarily the use of the energy concept in the solution of mechanical problems. This is the celebrated principle of Hamilton. It will serve very usefully as an introduction to what is usually referred to as advanced mechanics and about which the reader of this book may well expect to learn a little since it is basic for theoretical physics in general. We shall state the principle as a postulate and then endeavor to illustrate its application.

Let us consider a dynamical system which may be composed of many particles and hence may require a large number of coördinates to specify its state, i.e., the position of every particle at every instant. The total number of coördinates needed for this specification is called the number of degrees of freedom of the system. Now associated with the system there will be a certain kinetic energy K and, if it is a conservative system, a potential energy V. The former will be a function of the velocities of the particles, the latter a function of their positional coördinates. Let us form the function

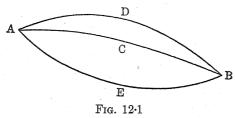
L = K - V. (12.1-1)

Hamilton's principle now states that the motion of the system of particles between any two instants t_0 and t_1 takes place in such a

way that the integral

$$\int_{t_0}^{t_1} L \ dt \tag{12.1-2}$$

has a value which is either less than its value for any other possible motion of the system in the interval from t_0 to t_1 or greater than its value for any other such possible motion. That is, the integral is either a maximum or a minimum or, as we may better say, has a stationary value with respect to other possible motions. We may perhaps render the matter somewhat clearer by using a pictorial representation. In Fig. 12·1 let us represent the state of the system at the instant t_0 by A and that at time t_1 by B. Let the path actually followed by the system be ACB, while ADB and AEB



represent two "varied" paths—possible paths for the system and performed in the same time. Now it must be remarked that our representation is very general. In certain simple cases of the plane motion of a single particle ACB, ADB, and AEB may represent possible geometrical orbits from point A to point B in the plane with the understanding that to be comparable they must of course all be traversed in the same time, viz., in the interval $t_1 - t_0$, and must correspond to the same initial and final states respectively. On the other hand, the representation can be purely schematic, i.e., A may merely represent for convenience the state of a system of several particles, etc. But in any case the meaning of the principle is that the time integral $\int_{t}^{t} L dt$ along the actual path traversed, ACB, is either greater than that along any other paths such as ADB and ACB or it is less than that along any such paths. It'is often convenient to state it as a minimum principle, i.e., that $\int_{-\infty}^{\infty} L dt$ is a *minimum* along the actual dynamical path. While in most problems encountered in practice this is correct, it

is not so in general. All we have a right to say is that it has a stationary value. The quantity L is termed the Lagrangian function or $kinetic\ potential$.

It must again be emphasized that we are here stating Hamilton's principle as a postulate, an assumption from which all special motions of bodies may be derived. As a first illustration of this view let us apply the principle to a simple problem. We shall take a single particle whose position is fixed by a single coördinate, for example, the distance along a straight line from some chosen origin. We are thus indeed restricting our attention to linear motion and hence our illustration is a very special one. The kinetic energy of the particle is then

 $K = \frac{1}{2}m\dot{x}^2, (12.1-3)$

if the particle is conceived to move along the x axis. Let us investigate the type of motion corresponding to a potential energy of the form

$$V = \frac{1}{2}kx^2, (12\cdot 1-4)$$

where k is a constant. The principle of Hamilton now demands that the integral

$$\int_{t_0}^{t_1} (\frac{1}{2}m\dot{x}^2 - \frac{1}{2}kx^2) dt, \qquad (12\cdot 1-5)$$

shall have a stationary value for the actual motion. This means that if we vary the true path slightly the corresponding variation of the integral will be zero. Of course, since in this case the geometrical shape of the path may not alter, a variation in the path signifies only an altered dependence of x on t, i.e., an altered mode of pursuing the linear path. The reader will remember that in the neighborhood of the maximum or minimum of a function of one independent variable t, the change in the function corresponding to a change dt in the variable is zero as far as first order quantities are concerned. If we denote the slight variation of the integral by $\delta \int_{t_0}^{t_0} \left(\frac{1}{2}m\dot{x}^2 - \frac{1}{2}kx^2\right) dt$, we have therefore the condition

$$\delta \int_{t_0}^{t_1} (\frac{1}{2}m\dot{x}^2 - \frac{1}{2}kx^2) dt = 0.$$
 (12·1-6)

Now the variation of the integral will clearly be the limit of the

sum of the variations in the integrand over all the infinitesimal time intervals between t_0 and t_1 . Hence we have

$$\int_{t_0}^{t_1} \delta(\frac{1}{2}m\dot{x}^2 - \frac{1}{2}kx^2) dt = 0.$$
 (12·1-7)

We are thus faced with the task of reducing $\delta(\hat{x}^2)$ and $\delta(x^2)$ to some common basis. One would naturally try to do this by expressing both in terms of δx . Let us consider the curves in Fig. 12-2. AB represents an element of the course of x as a function of t. CD is a portion of the varied curve. At A, the variation in x is δx . It is seen that this is for

is δx . It is seen that this is for a definite value of t. At B, the value of x is

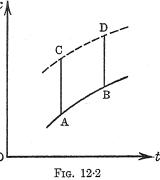
$$x + \dot{x} dt$$

and hence we may represent BD as either

$$\delta(x + \dot{x} dt), \quad (12\cdot 1-8)$$

or

$$\delta x + \frac{d}{dt} (\delta x) dt$$
. (12·1–9)



Consequently (12·1-8) and (12·1-9) are equal and hence

$$\delta(\dot{x} dt) = \delta(dx) = \frac{d}{dt} (\delta x) dt = d(\delta x). \qquad (12 \cdot 1 - 10)$$

Therefore the operations of δ and d are interchangeable. We can now write

$$\delta(x^2) = 2x \,\delta x, \qquad (12\cdot 1-11)$$

and

$$\delta(\dot{x}^2) = 2x \,\delta\dot{x} = 2\dot{x} \,\frac{d}{dt} \,(\delta x). \tag{12-1-12}$$

Eq. $(12\cdot1-7)$ then becomes

$$\int_{t_0}^{t_1} (m\dot{x}\,\delta\dot{x} - kx\,\delta x)\,dt = 0,$$

or using eq. $(12\cdot 1-12)$,

$$\int_{t_0}^{t_1} \left[m\dot{x} \, \frac{d}{dt} \, (\delta x) \, - \, kx \, \delta x \, \right] dt = 0. \tag{12.1-13}$$

Let us now evaluate the integral

$$\int_{t_0}^{t_1} m\dot{x} \, \frac{d}{dt} \, (\delta x) \, dt$$

by integration by parts. Letting $u = m\dot{x}$ and dv = d/dt (δx) dt, we have $du = m\ddot{x} dt$ and $v = \delta x$, so that

$$\int_{t_0}^{t_1} m\dot{x} \, \frac{d}{dt} \left(\delta x \right) \, dt \, = \, m\dot{x} \, \delta x \bigg]_{t_0}^{t_1} - \, \int_{t_0}^{t_1} m\ddot{x} \, \delta x \, dt.$$

But the two paths must have the same end points, hence $\delta x = 0$ for $t = t_0$ and $t = t_1$. Therefore eq. (12·1-13) becomes

$$\int_{t_0}^{t_1} (m\ddot{x} + kx) \, \delta x \, dt = 0. \tag{12.1-14}$$

It must be emphasized that this result holds for all δx , i.e., δx is perfectly arbitrary except that it be small and equal to zero for $t=t_0$ and t_1 . Hence the only way for the integral to be zero for all δx is for us to have

$$m\ddot{x} + kx = 0. \tag{12.1-15}$$

This then appears as the differential equation of the motion, and we see at once (Sec. 2.2) that the latter is simple harmonic with frequency

$$\nu = \frac{1}{2\pi} \sqrt{\frac{k}{m}}$$
 (12·1–16)

Hence from the assumption of Hamilton's principle together with the knowledge of the potential energy function we have been able to derive the equation of motion independently of any other dynamical assumption except that the kinetic energy shall be of the form $\frac{1}{2}m\dot{x}^2$. In other words we have made no essential use of the Newtonian laws of motion. Hamilton's principle thus appears as an independent basis for the development of mechanics. A special case indeed is no proof of this statement. We have merely tried to suggest its validity without proving it. Nevertheless the general proof has been carried out and may be found in more

advanced treatises. The power of the principle is more manifest in the treatment of the motion of systems with more than one degree of freedom. We shall discuss this with some detail in the next section.

12.2. Generalized Coördinates and Lagrange's Equations. an illustration of a more general application of Hamilton's principle, let us consider a mechanical system composed of a number of particles. If these particles are totally unconnected the total number of degrees of freedom or number of coördinates needed to specify the system will be three times the number of the particles. However if the particles are connected or otherwise constrained to move in certain ways (see Chapter VIII), the number of degrees of freedom will be reduced. Suppose the system in question has n. Now it should be clear that the choice of these n coordinates is possible in a very great variety of ways. They may of course be rectangular coördinates or certain combinations of these, or they may contain angles as well as lengths. In order to make our application most general we shall not specify the exact nature of the coördinates, but merely write them q_1, q_2, \ldots, q_n . All our description of the behavior of the system will then be in terms of the q's. In the last analysis, of course, the latter must be translated into some definite known set, which appear particularly appropriate for the description on the ground of simplicity. coördinates q_1, \ldots, q_n have been termed generalized coördinates. To express the kinetic energy in terms of them we note that

$$K = \sum_{i=1}^{1} m_i (\dot{x}_i^2 + \dot{y}_i^2 + \dot{z}_i^2), \qquad (12.2-1)$$

the summation being extended over *all* the particles. Let us suppose that the equations defining the generalized coördinates are expressed in the form

$$x_i = x_i(q_1, \dots, q_n), y_i = y_i(q_1, \dots, q_n),$$

 $z_i = z_i(q_1, \dots, q_n),$ (12.2-2)

for any particular i. We therefore have

$$\dot{x}_i = \frac{\partial x_i}{\partial q_1} \dot{q}_1 + \dots + \frac{\partial x_i}{\partial q_n} \dot{q}_n = \sum_{j=1}^n \frac{\partial x_i}{\partial q_j} \dot{q}_j, \qquad (12.2-3)$$

¹ Cf. Lindsay and Margenau, "Foundations of Physics" (Wiley, 1936), p. 131 for a deduction of Hamilton's principle from D'Alembert's principle.

with similar expressions for \dot{y}_i and \dot{z}_i . Substituting now into the expression for K, we get finally

$$K = \sum_{ki} a_{ki} \, \dot{q}_k \, \dot{q}_i, \qquad (12 \cdot 2 - 4)$$

where the summation is extended over both k and j running from 1 to n, and the a_{kj} are (in general) functions of the q_1, \ldots, q_n . The kinetic energy thus appears as a homogeneous quadratic function of the time derivatives of the generalized coördinates, or as we may call them the generalized velocities.

Now how does it stand with the potential energy? For a conservative system V is a function of the position coördinates of the particles. Hence it will be a function of the q_1, \ldots, q_n . Let us next apply the principle of Hamilton to such a system of n degrees of freedom. We express the stationary character of the time integral of the Lagrangian function for the actual motion as follows:

$$\delta \int_{t_0}^{t_1} (T - V) dt = \int_{t_0}^{t_1} \sum_{i} \left[\frac{\partial L}{\partial q_i} \delta q_i + \frac{\partial L}{\partial \dot{q}_i} \delta \dot{q}_i \right] dt = 0, \quad (12.2-5)$$

employing (12·1-1). Now since $\delta \dot{q}_i = \frac{d}{dt} (\delta q_i)$ from Sec. 12·1, we may integrate by parts

$$\int_{t_0}^{t_1} \sum \frac{\partial L}{\partial \dot{q}_i} \delta \dot{q}_i \, dt,$$

precisely as in the special problem in Sec. 12·1. The result is

$$\int_{t_0}^{t_1} \sum \frac{\partial L}{\partial \dot{q}_i} \delta \dot{q}_i dt = \left[\sum \frac{\partial L}{\partial \dot{q}_i} \delta q_i \right]_{t_0}^{t_1} - \int_{t_0}^{t_1} \sum \frac{d}{dt} \left\{ \frac{\partial L}{\partial \dot{q}_i} \right\} \delta q_i dt
= - \int_{t_0}^{t_1} \sum \frac{d}{dt} \left\{ \frac{\partial L}{\partial \dot{q}_i} \right\} \delta q_i dt,$$
(12.2-6)

since the part integrated out vanishes at the two limits because $\delta q_i = 0$ at both limits. This merely means that all the paths of the system have the same termini. Hence the statement of the principle becomes

$$\int_{t_0}^{t_1} \sum \left\{ \left[\frac{\partial L}{\partial q_i} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) \right] \delta q_i \right\} dt = 0.$$
 (12·2–7)

Now the n variations δq_i are arbitrary (the q's are independent)

 $(12 \cdot 2 - 9)$

and hence in order that the integral may vanish each separate bracket in the sum must vanish. This however gives us a set of n equations, viz.,

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = 0, \qquad (12.2-8)$$

as i runs from $1 \dots n$. These are the celebrated equations of Lagrange. We see that they are differential equations of the second order and that their number is precisely equal to the number of degrees of freedom of the system. These equations form a very general description of the motion of the system, and their solution leads to the expression of each q_i as a function of the time. There will of course be 2n arbitrary constants of integration involved; their evaluation depends on the boundary conditions of the system.

Let us note incidentally that for the case of a system of one degree of freedom where $q_i(=q)$ becomes simply the geometrical displacement and \dot{q} is the velocity, with $K = \frac{1}{2}m\dot{q}^2$, the single Lagrangian equation becomes the usual equation of motion, viz.,

$$\frac{d}{dt}(m\dot{q}) + \frac{\partial V}{\partial q} = 0,$$

$$m\ddot{q} = -\frac{\partial V}{\partial q},$$
(13)

or

where $-\frac{\partial V}{\partial q}$ appears as the force acting on the system.

Examination discloses that the Lagrangian equations (12·2-8) can be written in a form reminiscent of Newton's second law, provided that we define

$$p_i = \frac{\partial L}{\partial \dot{q}_i} \tag{12.2-10}$$

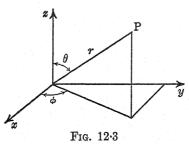
as the generalized momentum associated with the coördinate q_i , and further call

$$\frac{\partial L}{\partial q_i}$$

the generalized force associated with coördinate q. Then er

(12·2-8) says that the time rate of change of generalized momentum is equal to the generalized force, which reminds us of eq. (1·7-5). The reader might then be inclined to wonder wherein the advantage of Lagrange's equations consists. This comes principally from the fact that the Lagrangian equations being expressed in generalized form are immune to change in form when coördinates are transformed from one system to another: once we have expressed L in any system of coördinates, use of (12·2-8) automatically produces the correct equations of motion for those coördinates. It is apparent that this is not true of the simple Newtonian equation that mass times acceleration equals force. From an esthetic point of view the Lagrangian equations may be considered to have the additional merit of employing energy as the fundamental concept in their makeup; they make no mention of force

12.3. Application of Lagrange's Equations. We shall continue our discussion of Lagrange's equations with the solution of certain problems employing them. As the first illustration let us consider the motion of a single mass particle in a central field of force. We have, of course, treated this problem in the usual mechanical way



in Sec. 3.6. Our purpose here is to show how the general equations for such motion in spherical coördinates can be written with great ease by the Lagrangian method. There is often a gain in utility through this procedure. It will be recalled that the spherical coördinates r, θ , ϕ are defined through the accompanying

figure (Fig. 12·3), the point P being on a sphere of radius r, and its longitude (with respect to Ox) and colatitude being, respectively, ϕ and θ . It is seen that the spherical coördinates are related to rectangulars by the relations

$$x = r \sin \theta \cos \phi,$$

 $y = r \sin \theta \sin \phi,$
 $z = r \cos \theta.$

The kinetic energy of a particle is then

$$K = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2)$$

= $\frac{m}{2}(\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta \cdot \dot{\phi}^2),$ (12·3-2)

and, if the field is central (Sec. 3-6), the potential energy V is a function of r only, so that the Lagrangian function becomes

$$L = \frac{m}{2} (\dot{r}^2 + r^2 \dot{\theta}^2 + r^2 \sin^2 \theta \cdot \dot{\phi}^2) - V(r), \quad (12.3-3)$$

the system being one with three degrees of freedom with $q_1 = r$, $q_2 = \theta$, and $q_3 = \phi$. The Lagrangian equations (12.2-8) are then

$$\frac{d}{dt}(m\dot{r}) + \frac{dV(r)}{dr} - mr(\dot{\theta}^2 + \sin^2\theta \cdot \dot{\phi}^2) = 0,$$

$$\frac{d}{dt}(mr^2\dot{\theta}) - mr^2\sin\theta\cos\theta \cdot \dot{\phi}^2 = 0,$$

$$\frac{d}{dt}(mr^2\sin^2\theta \cdot \dot{\phi}) = 0.$$
(12·3-4)

From these equations the whole theory of central motion can be followed out. It will be observed that the longitude ϕ does not appear explicitly in any of the equations (12·3–4). This means that the generalized momentum component associated with ϕ is constant. Or we can say that $\dot{\phi}$ is determined from the third equation in terms of r and θ , viz.,

$$\dot{\phi} = \frac{C}{mr^2 \sin^2 \theta},\tag{12.3-5}$$

where C is a constant. This may be substituted into the first two equations yielding

$$\frac{d}{dt}(m\dot{r}) + \frac{dV}{dr} - mr\left(\dot{\theta}^2 + \frac{C^2}{m^2r^4\sin^2\theta}\right) = 0,$$

$$\frac{d}{dt}(mr^2\dot{\theta}) - \frac{C^2\cos\theta}{mr^2\sin^3\theta} = 0.$$
(12.3-6)

It is thus seen that the problem is effectively reduced to one with two degrees of freedom. We have already noted in Sec. 3.6 that

central field motion takes place in a plane. It being most simple to take this plane as the plane $\phi = 0$ (i.e., the xz plane), we have $\dot{\phi} = 0$ (i.e., C = 0, above) and consequently may write

$$\frac{d}{dt}(m\dot{r}) + \frac{dV}{dr} - mr\dot{\theta}^2 = 0,$$

$$\frac{d}{dt}(mr^2\dot{\theta}) = 0.$$
(12·3-7)

The law of areas, viz.,

$$mr^2\dot{\theta} = \text{const.},$$
 (12.3–8)

follows at once from the second equation, and if we substitute this into the first we obtain ultimately, on substitution of r = 1/u,

 the eq. (3.6-15) which we used in our previous discussion of central field motion.

A somewhat more elaborate illustration of the use of Lagrange's equations is provided by the following problem. Two small metal spheres A and B of equal mass m are suspended from two points A' and B' at a distance a apart on the same horizontal line by insulating threads of equal length l (Fig. 12-4). They are then charged with equal charges +e. If A is pulled aside a short distance q_1 and B a short distance q_2 (both small compared with a) in the plane formed by the equilibrium positions of both

threads, we are to determine the resulting motion in this plane. This is a system with two degrees of freedom. For the kinetic energy we have

$$K = \frac{1}{2}m(\dot{q}_1^2 + \dot{q}_2^2). \tag{12.3-9}$$

The potential energy will be made up of two parts, viz., the gravitational and the electrostatic. The former is simply the sum of that for each sphere considered as the bob of a simple pendulum. Hence (Secs. $4\cdot1$ and $8\cdot1$)

$$V_g = \frac{1}{2} \frac{mg}{l} (q_{1}^2 + q_{2}^2).$$
 (12·3–10)

The student will recall that from Coulomb's law the electrostatic

force of repulsion between the two point charges when separated by distance a is e^2/a^2 . The potential energy for this force when the charged balls are displaced will then be

$$V_e = \frac{e^2}{a + q_2 - q_1} \cdot \tag{12.3-11}$$

Hence for the Lagrangian function we have

$$L = \frac{1}{2} m (\dot{q}_1^2 + \dot{q}_2^2) - \frac{e^2}{a + q_2 - q_1} - \frac{1}{2} \frac{mg}{l} (q_1^2 + q_2^2), \quad (12.3-12)$$

to the approximation which considers the displacements as small. We are now ready to write the Lagrangian equations for the motion. Thus

$$\frac{d}{dt}(m\dot{q}_1) + \frac{mg}{l}q_1 + \frac{e^2}{(a+q_2-q_1)^2} = 0, \quad (12\cdot3-13)$$

$$\frac{d}{dt} (m\dot{q}_2) + \frac{mg}{l} q_2 - \frac{e^2}{(a + q_2 - q_1)^2} = 0.$$
 (12.3–14)

To solve, let us add the two equations. There results

$$\frac{d^2}{dt^2}(q_1+q_2)=-\frac{g}{l}(q_1+q_2). \qquad (12\cdot 3-15)$$

On the other hand if we subtract the two equations we get ultimately

$$\frac{d^2}{dt^2} (q_2 - q_1) + \frac{g}{l} (q_2 - q_1) - \frac{2e^2}{m(a + q_2 - q_1)^2} = 0. \quad (12-3-16)$$

Since both q_1 and q_2 are supposed small compared with a, we may expand the last term on the left in $(12\cdot3-16)$ writing

$$\frac{d^2}{dt^2} (q_2 - q_1) + (q_2 - q_1) \left(\frac{g}{l} + \frac{4e^2}{ma^3} \right) - \frac{2e^2}{ma^2} = 0. \quad (12 \cdot 3 - 17)$$

The solution of (12·3-15) is clearly

$$q_1 + q_2 = A_1 \cos(2\pi\nu_1 t + B_1),$$
 (12·3-18)

where

$$\nu_1 = \frac{1}{2\pi} \sqrt{\frac{g}{l}} \, \cdot \tag{12.3-19}$$

The solution of $(12\cdot3-17)$ is

$$q_2 - q_1 - \frac{2e^2}{ma^2\left(\frac{g}{l} + \frac{4e^2}{ma^3}\right)} = A_2 \cos(2\pi\nu_2 t + B_2), \quad (12\cdot3-20)$$

where

$$\nu_2 = \frac{1}{2\pi} \sqrt{\frac{g}{l} + \frac{4e^2}{ma^3}} \cdot \tag{12.3-21}$$

e

We finally have for the two displacements

$$q_{1} = \frac{-e^{2}}{ma^{2} \left(\frac{g}{l} + \frac{4e^{2}}{ma^{3}}\right)} + \frac{A_{1}}{2} \cos (2\pi\nu_{1}t + B_{1})$$

$$-\frac{A_{2}}{2} \cos (2\pi\nu_{2}t + B_{2}), \quad (12\cdot3-22)$$

$$q_{2} = \frac{e^{2}}{ma^{2} \left(\frac{g}{l} + \frac{4e^{2}}{ma^{3}}\right)} + \frac{A_{1}}{2} \cos (2\pi\nu_{1}t + B_{1})$$

$$+\frac{A_{2}}{2} \cos (2\pi\nu_{2}t + B_{2}). \quad (12\cdot3-23)$$

It may be noted that by shifting the origin for each ball the distance

$$\frac{e^2}{ma^2\left(rac{g}{l} + rac{4e^2}{ma^3}
ight)}$$
, (12·3–24)

the constant term may be made to disappear. In this case the motion of each ball appears as the composition of two simple harmonic motions with amplitudes and phases (A_1, A_2, B_1, B_2) which depend on the boundary conditions.

As a final illustration of the use of Lagrange's equations consider the situation depicted in Fig. 12.5.

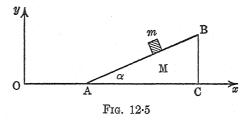
Here a wedge with projection ABC on the xy plane and with mass M is able to slide freely on the horizontal plane. The perfectly smooth slant surface AB of the wedge makes the angle α with the horizontal. On this surface a particle of mass m is free to move in the xy plane. The problem is to determine the motion of m

as well as the wedge. Referring to fixed axes as in the figure, we have for the kinetic and potential energies of the total system, respectively,

$$K = \frac{1}{2}M\dot{x}^2 + \frac{1}{2}m(\dot{x} + \dot{y}/\tan\alpha)^2 + \frac{1}{2}m\dot{y}^2, \quad (12\cdot3-25)$$

$$V = mgy. \quad (12\cdot3-26)$$

Here x is the distance OA, while y is the height of m above the x axis. The problem is one in two degrees of freedom. Note that we might equally well have chosen the x coördinate of m along with y as the second coördinate, i.e., $x + y/\tan \alpha$. The Lagrangian



method does not care what coördinates are used, provided that the number is correct for the nature of the system being studied, and provided that the Lagrangian function is correctly expressed in terms of the coördinates chosen. With the present choice we then have

$$L = \frac{1}{2}M\dot{x}^2 + \frac{1}{2}m\dot{y}^2 + \frac{1}{2}m(\dot{x} + \dot{y}/\tan\alpha)^2 - mgy.$$
 (12·3-27)

The Lagrangian equations become at once

$$\begin{split} M\ddot{x} + m\ddot{x} + m\ddot{y}/\tan\alpha &= 0, \\ m\ddot{y} + m\ddot{x}/\tan\alpha + m\ddot{y}/\tan^2\alpha + mg &= 0. \end{split} \tag{12.3-28}$$

The solution of these equations for x and y is left as an exercise. Note the interesting limiting case of $m/M \to 0$.

12.4. Wave Mechanics. We shall conclude the book with a brief discussion of the relatively recent application of mechanics to the study of atomic particles and atomic structure. This is known as wave mechanics. It endeavors to apply to the motion of particles some of the ideas associated with waves.

We have already seen (Sec. 3.9) how the Bohr-Rutherford theory considers a neutral atom to be composed of a nucleus charged

with positive electricity about which move electrons equal in number to the number of positive charges on the nucleus. electrons move in orbits according to the laws of classical mechanics. but with the significant difference that not all mechanically possible paths are allowed, but only those for which certain so-called quantum conditions are satisfied (recall particularly eqs. 3.9-1). latter serve to fix the possible orbital dimensions and energy values of the nucleus-electron system. (Recall also the corresponding problem for the electron in simple harmonic motion: Sec. 9.7.) Now it is important to note that in this theory the electron is still considered as a mass particle whose motion is to be treated largely as a classical mechanical problem. L. De Broglie has however suggested a quite different method of approach to this problem. He begins by assuming that an electron is not an entity localized in space but a periodic phenomenon extending throughout all space. More specifically, let us consider a free electron, that is, one not subject to any external force. De Broglie supposes that in the inertial system with respect to which the electron is at rest it is to be represented by an expression of the form

$$u(x_0, y_0, z_0, t_0) = f(x_0, y_0, z_0)e^{2\pi i \nu_0 t_0}.$$
 (12.4-1)

In this expression we are using x_0, y_0, z_0, t_0 to denote position and time in the system in which the electron is at rest. The righthand side represents a simple harmonic pulsation with frequency v_0 and amplitude $f(x_0, y_0, z_0)$, the phase being the same at all points in space. Both v_0 and $f(x_0, y_0, z_0)$ must now be looked upon as something fundamental for the physical significance of the particle. At first sight this point of view appears rather barren of meaning. However, let us next refer the electron to a reference system with respect to which it is moving with constant velocity v, say for convenience along the z axis. What will be the mathematical representation of the electron in this new set of axes which may be regarded as fixed from the standpoint of a hypothetical observer of the electron? Now in the special theory of relativity Einstein showed that when a physical event is represented in one set of axes by x_0, y_0, z_0, t_0 (where t_0 is time indicated by a clock fastened to this system), the same physical event will be represented in another set of axes moving with respect to the first with constant velocity

¹ The use of the complex form is for purposes of convenience only. Either real or imaginary part could be used equally well.

v along their common z axes by the quantities x, y, z, t where

$$x_{0} = x,$$

$$y_{0} = y,$$

$$z_{0} = \frac{z - vt}{\sqrt{1 - \frac{v^{2}}{c^{2}}}}, \quad t_{0} = \frac{t - \frac{v}{c^{2}}z}{\sqrt{1 - \frac{v^{2}}{c^{2}}}},$$

$$(12.4-2)$$

the quantity c being the velocity of light. These are the famous Lorentz-Einstein transformation equations. We shall not stop here to derive them, but they are so important for modern theoretical physics that the student should familiarize himself with them and their meaning without delay. We note only that they are a kind of generalization of the Newtonian transformation equations discussed in Sec. 1·10 and effectually reduce to them when $\frac{v}{c} \ll 1$, i.e., when the relative velocity of the two systems is small compared with the velocity of light. This is, of course, the situation throughout classical mechanics, but in the case of atomic particles the radical $\sqrt{1-\frac{v^2}{c^2}}$ can become very different from unity.

To solve the problem of De Broglie we must now substitute from $(12\cdot4-2)$ into $(12\cdot4-1)$. However, in so doing we must not overlook the fact that ν_0 , being a frequency, depends on the time and must also undergo transformation. Thus, if we concentrate on a particular value of z, corresponding intervals of time in the two inertial systems from the last of eqs. $(12\cdot4-2)$ will be given by

$$dt_0 = \frac{dt}{\sqrt{1 - \frac{v^2}{c^2}}},$$

¹ See, for example, Page, "Theoretical Physics," p. 452. These equations and their meaning are now quite commonly discussed in the more recent elementary textbooks of physics. Thus see C. W. Miller, "Introduction to Physical Science" (John Wiley, New York, 1932), Chap. XXXIII; A. E. Caswell, "An Outline of Physics" (Macmillan, New York, 1928, Chap. XXXIX). See also the recent interesting intermediate text "An Outline of Atomic Physics," by members of the physics staff at the University of Pittsburgh (Wiley, New York, 1933), Chap. XIII. This is an excellent elementary treatment. Further consult G. E. M. Jauncey, "Modern Physics," Chap. XIX.

and therefore ν_0 , which is proportional to $1/dt_0$, will transform to ν , where

$$\nu = \frac{\nu_0}{\sqrt{1 - \frac{v^2}{c^2}}}.$$
 (12·4–3)

On finally carrying out the indicated transformation we get for the representation of the electron in the system with respect to which it moves with constant velocity v

$$u'(x, y, z, t) = f\left(x, y, \frac{z - vt}{\sqrt{1 - v^2/c^2}}\right) e^{2\pi i v \left(t - \frac{v}{c^2}z\right)}.$$
 (12·4-4)

We recognize this at once (Sec. 10.4) as a plane progressive harmonic wave of frequency ν moving in the positive z direction with velocity

$$V = \frac{c^2}{v} \cdot \tag{12.4-5}$$

The amplitude f is, of course, variable in space and time. the inertial system with respect to which it is moving with constant velocity v. the electron appears to be equivalent to a harmonic wave. It is interesting to note from (12.4-5) that the velocity of this wave is never less than the velocity of light, since $v \leq c$ for a particle, a fact that is one of the consequences of the special theory The fact that $V \ge c$ might seem to detract from the of relativity. value of De Broglie's viewpoint. However, he was able to show that if one considers the particle to be represented not by a single wave but by a whole group of harmonic waves with slightly different frequencies clustered about an average frequency ν and moving in the same direction with slightly different velocities, the velocity of the point of phase agreement of such a group, i.e., the so-called group velocity, is precisely equal to the particle velocity v. The reader should work this out for himself. (See Problem 12 at the end of the chapter.)

De Broglie further assumed that there is a fundamental relation between the *energy* of the electron considered as a particle moving with respect to the fixed axes and its frequency considered as a wave. This relation is the famous quantum theory equation

$$E = h\nu, \qquad (12.4-6)$$

where h is the universal constant of Planck (Secs. 3.9 and 9.7).

Now on the theory of relativity the energy of a freely moving particle is given by the simple relation

$$E = mc^2, (12.4-7)$$

where m is the mass of the particle with respect to the system in which it is moving with velocity v. It can then be shown that

$$m = \frac{m_0}{\sqrt{1 - \frac{v^2}{c^2}}},\tag{12.4-8}$$

 m_0 being the so-called *rest* mass of the particle, i.e., its mass in the system in which it is at rest or which moves with it.¹ The combination of eqs. $(12\cdot4-5)$, $(12\cdot4-6)$, and $(12\cdot4-7)$ at once yields the important relation for the wave length λ of the wave representing the electron. Thus

$$\lambda = \frac{h}{mv} = \frac{h}{p},\tag{12.4-9}$$

if by p we represent the *momentum* of the electron with respect to the fixed system. Experiments in which electrons are reflected from the surfaces of single metallic crystals have verified eq. $(12\cdot4-9)$ and the fact therefore of the wave nature of the electron.²

The considerations just presented refer primarily to a free electron. The problem of describing in terms of the wave point of view an electron acted on by a force is more complicated. Nevertheless Schrödinger and others have shown how to treat for example the case of an electron moving in simple harmonic motion and an electron moving about a positive nucleus as in the atom of hydrogen, as well as more elaborate cases. The result of the analysis, as in the analogous Bohr method, is to fix the possible energy values of the system in question, i.e., to quantize it. We may get some idea as to how this is done by discussing briefly the problem of the electron in simple harmonic motion. It will be recalled that in Sec. 9.7 we carried through the quantization of this

¹ See again Page, loc. cit., p. 463. Or "Outline of Atomic Physics," loc. cit., p. 265. O? Jauncey, "Modern Physics," p. 455.

² For further elementary discussion of this fascinating development of modern physics see "Outline of Atomic Physics," *loc. cit.*, p. 142. Also Jauncey, "Modern Physics." The more advanced treatment may be found in Ruark and Urey, "Atoms, Molecules and Quanta."

motion using the Bohr quantum condition. It will be very instructive to compare this with the Schrödinger wave mechanics method.

In every problem in wave mechanics it is first necessary to find the wave equation that corresponds to the particular dynamical system being studied. To treat this problem in a general way is of course outside the scope of the present text but it will doubtless do no harm if we follow for the present illustration a formal, if apparently arbitrary, procedure which the reader will find justified in treatises on the subject. It will be recalled that the energy equation for simple harmonic motion (Sec. 4·1) is

$$\frac{1}{2}mv^2 + \frac{1}{2}kx^2 = E, (12.4-10)$$

where k is the so-called "stiffness" or "elastic" constant and E is the total energy. From eq. (2·2–13) we have

$$k = 4\pi^2 \nu_f^2 m, \qquad (12.4-11)$$

if v_f is the frequency of the motion. Let us write $(12\cdot4-10)$ in terms of the momentum p=mv. It then becomes $(12\cdot4-10)$

$$\frac{p^2}{2m} + \frac{kx^2}{2} = E. (12.4-12)$$

The meaning of this equation from the standpoint of classical mechanics is sufficiently clear and need not be further emphasized. Wave mechanics proceeds to give it a new interpretation by assigning a new operational significance to the symbol p. Thus instead of interpreting it in the usual mechanical sense, let us say that p is equivalent to the differential operator $\frac{h}{2\pi i} \frac{d}{dx}$. That is, if $\psi(x)$ is any function of x, we shall assume the identity

$$p\psi(x) = \frac{h}{2\pi i} \frac{d\psi}{dx} \cdot \tag{12.4-13}$$

This will further be extended to mean that

$$p^2 \psi(x) = \left(\frac{h}{2\pi i}\right)^2 \frac{d^2 \psi}{dx^2}$$
 (12.4-14)

 1 It will be recalled that we have already found the operator notion valuable in Sec. 9.7 in our discussion of vibrating systems.

Our new interpretation of eq. (12·4-12) then is this: if we operate on any function $\psi(x)$ with the operator $\frac{1}{2m}\left(\frac{h}{2\pi i}\right)^2\frac{d^2}{dx^2}+\frac{kx^2}{2}$, the result is the same as multiplying $\psi(x)$ by the constant E. Written down in formal terms, however, this gives us at once the second order differential equation

$$\frac{d^2\psi}{dx^2} + \frac{8\pi^2 m}{h^2} (E - \frac{1}{2}kx^2)\psi = 0.$$
 (12.4–15)

Comparing this equation with eq. $(10\cdot4-3)$, for example, we see that it may be interpreted as the wave equation for plane harmonic waves along the x axis with variable wave length

$$\lambda = \frac{h}{\sqrt{2m(E - \frac{1}{2}kx^2)}} \cdot \tag{12.4-16}$$

The eq. (12.4-15) is then taken as the wave equation (the socalled Schrödinger wave equation) associated with the simple harmonic oscillator. Now that we have our wave equation, what are we going to do with it? In the case of the wave equation for a finite stretched string discussed earlier in this chapter we sought for a solution giving the displacement of the string as a function of x and t subject to certain boundary conditions. Now this is precisely what Schrödinger does with the eq. (12.4-15). He looks for a solution giving ψ as a function of x (it is already assumed to be harmonic in time) which satisfies the general boundary condition that it shall be finite, continuous and single-valued for all x and shall vanish at infinity. The solution of this type of problem has long been familiar to mathematicians under the general name of boundary value problem. It is a little too lengthy for us to embark on here. It will be sufficient to say, however, that the analysis shows that it is impossible to obtain solutions of (12.4-15) which satisfy this boundary condition for all values of the energy E. In fact it develops that such solutions are obtained only for values of the energy of the following form

$$E_n = (n + \frac{1}{2})h\nu_f, \qquad (12.4-17)$$

where $n = 0, 1, 2, \ldots$ These are then called the *eigen-values* or *characteristic values* of the simple harmonic oscillator. Physically speaking they are the allowed energy values of the oscillator and

correspond to the quantization of the motion. It is very interesting to compare the result of the Schrödinger quantization with that of the Bohr quantum condition method — (eq. 9.7-10). It is seen that the two are alike save that $n+\frac{1}{2}$ in the former replaces n in the latter. Experimental applications indicate that (12.4-17) is more in accord with observed facts than (9.7-10). It is however important to note that the quantization process in both cases results from the imposition of a boundary condition: in the Bohr theory this is the quantum condition applied to the classical mechanical motion, while in the Schrödinger wave mechanics it is a condition analogous to those imposed on the wave equation representing a vibrating string.

The true physical meaning of the Schrödinger wave equation can be appreciated only after a comprehensive study of the foundations of wave mechanics which is really a new method of attacking physical problems in the atomic domain. All the problems worked out by the Bohr theory (such as the hydrogen atom and atoms of more complicated structure) have been solved also with wave mechanics and the results have been remarkably successful. Our aim in this section has been merely to provide a brief and somewhat formal introduction to this new discipline and in particular to indicate what an essential role a clear understanding of classical mechanics plays in its elucidation.¹

PROBLEMS

- 1. Use Hamilton's principle to determine the path of a particle which moves freely in three-dimensional space, i.e., under the action of no forces.
- 2. Use Hamilton's principle to derive the equation of motion of a particle moving in a straight line with potential energy = K/x, where x is the distance of the particle from some arbitrarily chosen origin.
- 3. Using Euler's theorem for homogeneous functions, show that for a dynamical system with n degrees of freedom (Cf. eq. 12-2-4)

$$\sum_{i=1}^{n} \dot{q}_i \frac{\partial K}{\partial \dot{q}_i} = 2K.$$

¹The modern version of wave mechanics is known as quantum mechanics. For further reading in this field the following books are recommended: Lindsay & Margenau, "Foundations of Physics" (Wiley, 1936), Chap. 9; G. E. M. Jauncey, "Modern Physics" (D. Van Nostrand, 3d ed., 1948), Chap. 17; L. Pauling and E. B. Wilson, "Introduction to Quantum Mechanics" (McGraw-Hill, 1935).

4. Consider a conservative system. Define

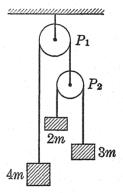
$$p_i = \frac{\partial K}{\partial \dot{q}_i}$$

as the generalized momentum associated with q_i . Taking the q_i 's in turn as the rectangular, spherical and cylindrical coördinates of a single particle, evaluate the corresponding p_i in each case and comment on its physical significance.

5. A particle of mass m moves along the x axis in simple harmonic motion. Calculate the time integral of the Lagrangian function over the actual dynamical path $x=A\sin\omega t$ from t=0 to t=T/4, where $T=2\pi/\omega$. Consider a varied path $x=A\sin\omega t+\lambda\sin4\pi t/T$, which has the same end points as the dynamical path at t=0 and t=T/4, if λ is a constant parameter. Find the time integral of the Lagrangian function over the varied path from t=0 to t=T/4 and compare with the previous result.

to t = 1/4 and compare with the previous result. Connect this with the fundamental physical significance of Hamilton's principle.

- 6. Use Lagrange's equations to determine the motion of the masses in the frictionless pulley set-up shown in the accompanying figure. The movable pulley P_2 has mass 2m.
- 7. Referring to Fig. 12.5, let the surface AB be completely rough and replace the particle m by a circular cylinder of mass m which can roll without slipping down the plane. Determine the motion of wedge and cylinder by the use of Lagrange's equations.



8. For a conservative system consider the expression

$$H = \sum_{i=1}^n p_i \dot{q}_i - L,$$

where L is the Lagrangian function and p_i is to be interpreted as in Problem 4. Show that H is a function of the p_i and q_i alone. Further show that H is constant in time and that it equals K+V, i.e., the total energy of the system. Finally show that

$$\frac{\partial H}{\partial q_i} = -\dot{p}_i, \quad \frac{\partial H}{\partial p_i} = q_i.$$

 $H(p_i,q_i)$ is called the *Hamiltonian* of the system, and the foregoing equations are called the *canonical* equations of motion, or the equations of motion in Hamiltonian form. They have proved particularly valuable in the application of mechanics to atomic structure problems.

9. Show that for the simple harmonic oscillator of mass m and stiffness constant k, the Hamiltonian is

$$H=\frac{p^2}{2m}+\frac{kq^2}{2}.$$

Obtain the canonical equations. Work the same problem for the central field motion in spherical coördinates.

- 10. According to Fermat's principle, the ray of light from any point A to any other point B in space can be determined by the assumption that the time f of travel from A to B is a minimum. Show that this leads to the experimentally observed laws of reflection and refraction of light at a plane surface. Show that Hamilton's principle applied to a light particle or photon of energy $h\nu$ is equivalent to Fermat's principle.
- 11. If a spherical light wave starts at time t=0 to move with velocity c from the origin of a fixed system of axes (x,y,z), at time t (as measured in this system) it will have reached the surface of a sphere whose equation is

$$x^2 + y^2 + z^2 - c^2 t^2 = 0.$$

Show that the equation of this sphere has the same form in the system of axes (x_0, y_0, z_0) moving with respect to the first with constant velocity v along their common z axes. (The origins of the two systems are assumed to pass each other at time $t = t_0 = 0$.) That is, using the Lorentz-Einstein transformation equations (12·4–2), show that the expression above becomes

$$x_0^2 + y_0^2 + z_0^2 - c^2 t_0^2 = 0.$$

Note that the theory of relativity supposes that c is a universal invariant. Comment on the physical significance of the result of the transformation.

12. Consider a group of plane harmonic De Broglie matter waves moving in the z direction with frequencies included in the range $\nu - \delta \nu$ to $\nu + \delta \nu$, where ν is a *mean* frequency and $\delta \nu \ll \nu$. Any particular wave in the group may be represented by

$$A(\nu + \epsilon) \cos 2\pi (\nu + \epsilon) \left[t - \frac{z}{V(\nu + \epsilon)} \right],$$

where $\nu + \epsilon$ is the frequency and $\epsilon \leq |\delta\nu|$. The velocity of the wave is $V(\nu + \epsilon)$ and is supposed to be a function of frequency. Consider particularly two waves in the group for which $\epsilon = \epsilon_1$ and ϵ_2 respectively. Show that if they agree in phase at some point z at a certain t, this point of phase agreement will move with velocity

$$U = \frac{1}{\frac{d\left(\frac{\nu}{V}\right)}{d\nu}},$$

the group velocity of the group of waves. Prove that for the case of De Broglie waves U = v, where v is the velocity of the particle represented by the wave group.



SUPPLEMENTARY PROBLEMS

1. Use the method of "dimensions" to derive the functional dependence of the period P of a simple pendulum on g and the length of the string. Discussion: Recall that with every mechanical quantity there is associated a dimensional formula involving the three fundamental dimensions of mass, length and time, represented symbolically by M, L, and T. Thus the dimensional formula for velocity is L/T or LT^{-1} ; that for acceleration is LT^{-2} ; that for force, MLT^{-2} , etc. To have physical meaning, every equation in mechanics must be consistent dimensionally; that is, both sides must have the same dimensions.

In the present example, if l = length of string, $m = \text{mass of bob and } \theta = \text{angular amplitude of swing, assume that } P$ is represented by the formula

$$P = l^x m^y g^z \theta^u.$$

P has the dimensions of T and hence the right-hand side must have the same dimensions. From this fact find the relations that x, y, z, u must satisfy in order to make the equation dimensionally consistent. Thus show that $P \propto \sqrt{l/g}$. The factor 2π in the actual formula is undetermined by the dimensional method. It is, of course, dimensionless. For a detailed account of dimensional analysis, see the interesting book by P. W. Bridgman "Dimensional Analysis" (Yale University Press, 2d ed., 1931).

- 2. Derive the functional form of Kepler's third law by the use of the method of dimensions. Do the same for Jurin's law for the rise of a liquid in a capillary tube (Sec. 11.8).
- 3. Show that when a mass particle moves in a circular path with a constant speed equal to that which it would gain in freely falling under gravity a distance equal to one-half the radius of the circle, the centripetal force on the particle is equal to its weight. This result is due to Huyghens (1629–1695) and is of considerable historical interest since it marks the beginning of the clear distinction between mass and weight.
- 4. A particle is constrained to fall under gravity from the highest point of a vertical circle along a chord of the circle. Prove that the time of descent is the same for *every* chord passing through the highest point. Prove also that the time of descent from all points on the circle is the same along the chords from those points passing through the lowest point of the circle.
- 5. Consider the parabola with equation $y = a kx^2$ with the xy plane vertical. A particle moves from rest under gravity from the focus. If it is constrained to move along some line extending from the focus to the parabola, find the equation of the line along which it will meet the parabola most quickly. Show that the length of this line is equal to that of the latus rectum.

- 6. A particle moves in a circle of radius R under the attraction of a force varying inversely as the square of the distance from the center of the circle. Use the expression for the centripetal force to derive Kepler's third law of planetary motion for this special (i.e., degenerate) case.
- 7. In the motion of a comet about the sun in a parabolic orbit, prove that the comet crosses the orbit of any given planet with a velocity $\sqrt{2}$ times that which the planet has at the same place (assuming that the orbit of the latter is approximately circular).
 - 8. The solution of Kepler's equation [Sec. 3.8, eq. (3.8-25)], viz.,

$$M = E - \epsilon \sin E,$$

where E is the eccentric anomaly, ϵ the eccentricity and M the mean anomaly, is of great importance, both in celestial mechanics and in the Bohr theory of atomic structure. The desired solution consists in finding E as a function of M and hence of t. This is equivalent to finding the radius vector r as a function of t. Indicate how a graphical solution may be obtained by plotting a sine curve and a straight line and getting their point of intersection. In this way construct graphical plots of E as a function of M for values of $\epsilon = 0.1$, 0.3, 0.5, 0.8.

- 9. A particle moves in an elliptical orbit in a central field of force varying inversely as the square of the distance from the force center (one focus of the ellipse). Calculate the time average of the radius vector in terms of the major axis and eccentricity of the orbit. Apply numerically to the case of the earth and sun.
- 10. A particle moves in an elliptical orbit in a central field of force varying directly as the distance from the force center (the center of the ellipse). Calculate the time average of the radius vector in terms of the constants of the orbit. What is the expression for the period of revolution in this type of motion? Mention a physical illustration of this motion. Cf. Sec. 3.5.
 - 11. A particle moves in a plane orbit with polar equation

$$r = \frac{\epsilon p}{1 - \epsilon \cos \gamma \theta},$$

where γ is a dimensionless constant very close to unity. Show that this orbit represents an ellipse of eccentricity ϵ and parameter p, whose perihelion slowly rotates. Find the rate of perihelion motion and make a sketch of the orbit. Assuming that the particle moves in a central force field, find the expression for the law of force, and in particular note the deviation from the inverse square law. All the planets show this perihelion motion to a certain extent, that of Mercury being the largest (574 seconds of arc per century). Classical Newtonian mechanics has been unable to account for the whole of this as being due to the perturbing influences of other planets. The theory of general relativity of Einstein has seemingly removed the difficulty (cf., e.g., G. E. M. Jauncey, "Modern Physics," p. 460). It is interesting to note that the application of special relativity mechanics to the motion of an electron about a nucleus in the Bohr theory leads to an orbit of precisely the form indicated

above with $\gamma = \sqrt{1 - 4\pi^2 e^4 Z^2/k^2 h^2 c^2}$, where e is the charge on the electron, Z is the number of positive charges on the nucleus, c is the velocity of light, h is Planck's constant, and k is any integer. See Ruark and Urey, "Atoms, Molecules and Quanta," p. 132.

- 12. A force $\mathbf{F} = \mathbf{i}F_0 + \mathbf{j}kte^{-at}$ acts on a particle, where F_0 is constant and equals 10/e dynes, $a = 1/20 \sec^{-1}$ and k = 200 dynes/sec. Find the momentum produced (direction and magnitude) in 1/10 second. Plot the magnitude of F as a function of t.
- 13. An alpha particle with initial velocity 2×10^9 cm/sec is shot *directly* at a nucleus of atomic number N = 80 (i.e., charge on the nucleus is 80 e):

(a) Assuming that the nucleus stays at rest, find the minimum distance of approach.

(b) What is the potential energy of the alpha particle at the minimum distance?

14. Find the way in which the central force must vary with the distance from the force center in order that a particle may describe the equiangular spiral

 $r = e^{a\theta}$

the force center being assumed to be at the pole (a is a constant). Use the law of areas to find the expression for the resultant velocity of the particle as a function of r. Solve the same problem for the spiral of Archimedes for which $r = a\theta$ (a is again a constant).

15. A particle of unit mass moves under the influence of no force save a resisting force of the form A + Bt where A and B are constants. If the particle comes to rest in time t_0 , find the total distance traversed and show that it is equal to $\int_0^{t_0} t(A + Bt) dt$. Try to generalize this result.

16. In a straight rod of length l, the line density varies linearly from ρ_0 at one end A to ρ_1 at the other end B. Find the mass of the rod and the position of the center of mass. Also find the gravitational force with which the rod attracts a unit mass at point P such that PA is perpendicular to the rod and equal to h. Through what point of the rod does this force pass?

- 17. Find the law of density distribution in a sphere so that its attraction for a particle on its surface shall be independent of the size of the sphere.
- 18. A particle of mass m moves in a straight line subject to an attractive force directed toward a fixed point on the line and varying inversely as the square of the distance from the point. It is assumed to start from rest at distance a from the fixed point. Discuss the motion and in particular find the possible energy values of the motion consistent with the Bohr quantum condition $\oint m\dot{x} dx = n\hbar$ (Sec. 3.9 and Sec. 4.4). These paths are the so-called "pendulum orbits" of the Bohr atomic theory. It should be shown that the

quantized energy values agree with those obtained for elliptical motion if an inverse square field for $n_1 + n_2 = n$. Physically speaking, these orbits are ruled out since they involve collision with the force center, i.e., the nucleus in the atomic problem.

19. An electron moves in a quantized elliptical orbit about a positive nucleus as in the Bohr theory of the hydrogen atom. Find the expression for the fraction of the whole period of revolution which the electron spends on the average in a spherical shell of thickness dr at distance r from the nucleus. Hint: Use the expression for the eccentric anomaly, Sec. 3.8. Hence show that if the electron's motion is assumed to be equivalent to a spherical distribution of negative charge about the nucleus the average fraction of the total

charge e in the spherical shell dr at distance r is $de = \frac{er dr}{\pi a^2 \sqrt{\epsilon^2 - 1 + \frac{2r}{a} - \frac{r^2}{a^2}}}$

Plot the distribution function. Integrate over r to get the total charge; this of course should be e.

20. Calculate the expression for the *potential* produced at any point distant r from the nucleus by the spherical distribution of the previous problem (Problem 19). Note: Electrical potential is calculated in the same way as gravitational potential. Care should be taken to distinguish between the potentials due to a given shell at points *inside* and *outside* the shell (Sec. 4.5). Use appropriate numerical values and plot the potential as a function of r.

21. Find the force acting on a particle of unit negative charge at point r in the field produced by a positive nucleus and the spherical distribution of negative charge described in Problem 19 above. This type of calculation has been of value in the Bohr theory of the structure of atoms with more than one electron. (Consult Ruark and Urey, p. 200.)

22. Use Gauss' law (Sec. 4.6) to show that Laplace's equation, viz.,

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} = 0$$

in spherical coördinates $(x=r\sin\theta\cos\phi,\,y=r\sin\theta\sin\phi,\,z=r\cos\theta)$ becomes

$$\frac{1}{r^2}\frac{\partial}{\partial r} \left(r^2 \, \frac{\partial V}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \, \frac{\partial V}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 V}{\partial \phi^2} = 0.$$

Hint: Apply the law to the volume element in spherical coördinates, viz., $r^2 \sin \theta \ d\theta \ dr$, whose surface consists of two spherical elements, two conical elements, and two plane elements.

23. A sphere of radius a contains electric charge distributed throughout its volume with constant density ρ . Use Gauss' law (Sec. 4-6) to obtain the magnitude of the electric intensity due to this charge distribution (i.e., the force on a unit charge) at any point distant r from the center of the sphere.

Thus show that for a point *inside* the sphere, the intensity has the magnitude $\frac{4\pi\rho r}{3}$, while *outside* it has the value $\frac{4\pi\rho a^3}{3r^2}$. Hint: Recall Coulomb's law of force for electric charges and compare with the gravitational analysis of Secs. 4.5 and 4.6.

24. In the preceding problem (Problem 23) calculate the electric potential (Sec. 4-3) for a point distant r from the center of the sphere both inside and outside. Thus obtain $V_i = -\frac{2}{3}\pi\rho(3a^2 - r^2)$ and $V_0 = -\frac{4\pi\rho a^3}{3r}$. Prove that these satisfy Poisson's and Laplace's equations respectively [eqs. (4-6-13) and (4-6-14)]. Hint: Use the latter expressed in spherical coördinates as in Problem 22 above. Work the same problem for the case of gravitational attraction.

25. A sphere of radius a contains electric charge distributed throughout its volume with density $\rho=k(a^2-r^2)$. Show that $k=\frac{15Q}{8\pi a^5}$, where Q is the total charge. Find the magnitude of the electric intensity at points inside and outside the sphere and do the same for the potential. Check by substitution into Poisson's equation.

26. Consider a conducting sphere of radius a. A point charge +e is placed at the point P distant u from the center of the sphere, O, where u > a. Prove that if a charge e' = -ae/u is placed at a point on OP inside the sphere distant $b = a^2/u$ from O, the potential at all points on the sphere will be zero. The charge e' is called the *electrical image* in the sphere of the charge e. The use of electrical images is very valuable in the solution of problems in electrostatics.

27. A boat with mass 2 tons (including that of the passenger) is moving through the water with an initially constant velocity of 20 ft/sec with respect to the shore. The passenger throws a 2-lb ball with an initial velocity of 30 ft/sec relative to the boat in the direction of motion of the boat. How much work does the man do in throwing the ball?

28. A gun and its carriage of mass M rest on a rough horizontal surface for which the coefficient of friction is μ . If the gun is fired horizontally so that the initial velocity of the projectile of mass m relative to the gun is V, find the distance which the carriage recoils.

29. A particle of mass m is acted on by a force $\mathbf{F} = kr^{-3}\mathbf{r}$ where k is a constant. How much work is done by the force when the particle moves in the xy plane in the arc of a circle from the position

$$r = ai$$

to the position

$$\mathbf{r}_2 = a\mathbf{i}$$
?

Obtain the answer in two different ways.

^{30.} Two forces act on a given particle. Prove that if the sum and difference of the two forces are perpendicular to each other, the two forces are of equal

- magnitude. Prove also that if the magnitudes of the sum and difference respectively of two forces are equal, the two forces are perpendicular. Hint: Use the scalar product of two vectors.
- 31. Two electrostatic point charges +2e and -e are at a distance a apart. Find the neutral point, i.e., that at which a unit charge would be in equilibrium under the action of the two given charges.
- 32. A particle is constrained to lie on the convex side of a smooth ellipse and is acted on by forces of magnitude F_1 and F_2 directed toward the two foci respectively, and a force of magnitude F_c directed toward the center. In whater position will the particle be in equilibrium?
- 33. A particle is placed at point P distant $r_1, r_2, \ldots r_n$ from the fixed points $O_1, O_2, \ldots O_n$. Forces act on the particle directed along $O_1P, O_2P, \ldots O_nP$ and are in magnitude proportional respectively to $r_1, r_2, \ldots r_n$. Use the principle of virtual work to show that the surface on which the particle will be in equilibrium in all positions is a sphere.
- 34. A flexible cord is suspended from two fixed points on the same horizontal line. What must be the law of variation of the mass per unit length in order that its shape may be a semicircle whose diameter is horizontal?
- 35. Prove that the resultant of several non-coplanar couples has a moment which is the resultant of the moments of the individual couples.
- 36. Find the center of mass of a sector of a plane uniform circular disk the arc of which subtends an angle of 30°. Also find its moment of inertia about an axis perpendicular to its plane and passing through the center of mass.
- 37. A plane ABC intersects the three coördinate axes in the points A, B, C respectively, forming the tetrahedron O-ABC. Four forces act normally to the faces of the tetrahedron along the three coördinate axes respectively and along the normal to ABC passing through O. These forces are proportional in magnitude to the areas of the faces to which they are respectively normal. Prove that they are in equilibrium. Generalize the result to the case of any tetrahedron. State the connection between the result of this problem and the laws of hydrostatics.
- 38. A uniform triangular plate with two sides equal to 10 cm and 15 cm respectively and the included angle 30° has a mass of 100 grams. It lies in a horizontal plane, suspended from a fixed point 25 cm above the plane by three strings attached to the corners. Find the tension in each of the strings. In particular find the relation between these tensions and the lengths of the strings.
- 39. A thin uniform rod has a length l and mass m. It is suspended by a vertical thread attached to its center O. A particle of mass M is placed at point P distant r from O so that OP makes the angle θ with the rod. What is the total force moment about the axis of suspension due to gravitational attraction?
- 40. A bar magnet of magnetic moment M (equal to ml where m is the strength of the poles, assumed to be concentrated at the ends of the magnet,

and δ is the length of the magnet) is suspended by a wire perpendicular to the lines of force of a uniform magnetic field of intensity H (that is, H dynes per unit pole or gauss). What torque or force moment is required to hold the magnet at an angle θ with respect to the direction of the field? How much work is done in the motion of the magnet from a position parallel to the field to a position making angle θ with the field?

If the magnet has moment of inertia *I* about a perpendicular axis through the center, find the frequency of the vibrations it will make if displaced slightly from its position of equilibrium with respect to the field and then let go.

- 41. Prove that the analytical conditions for the equilibrium of a rigid body retain the same form if oblique axes are used instead of rectangular axes.
- 42. Use the theorem of Pappus on centroids to find the expression for the volume of a torus or anchor ring.
- 43. A uniform brass cylinder of radius 1 cm and length 50 cm is allowed to swing in a vertical plane about one end through an angle of 30° each side of the plumb line. Find the period of its oscillations.
- 44. A heavy block of wood is suspended by a stiff but massless vertical wire. A bullet is fired into the block. Indicate how the velocity of the bullet may be obtained. Derive the formula

$$v = \frac{2(M+m)k\sqrt{gh}\sin\theta/2}{mp}$$

where: m= mass of the bullet; M= mass of the block; k= radius of gyration of the block with the bullet in it; v= velocity of the bullet; p= the distance of the line of motion of the bullet from the point of suspension of the block; h= the distance of the center of mass of the block from the point of suspension; $\theta=$ the angle through which the block is deviated from the vertical. This arrangement is known as a ballistic pendulum.

- 45. A uniform thin hemispherical shell is constrained to move with its pole fixed. Find the equation of the momental ellipsoid.
- 46. A particle is projected with velocity V on a smooth horizontal plane. Show that because of the rotation of the earth the particle will describe an arc of a circle of radius $\frac{V}{2\omega \sin \lambda}$, where ω is the angular velocity of the earth and λ is the latitude. Work out a numerical example.
- 47. The contour of an elliptical plate of unit thickness has the equation $Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0$. Find the moment of inertia of the plate about a diameter parallel to the x axis.
- 48. A thin circular sheet of radius r and mass m falls from rest in a vertical position to a horizontal floor. If the bottom in contact with the floor does not slip, with what kinetic energy will it strike and what will be the velocity of the top end?

49. What is the kinetic energy of rotation of a right circular cone with height 1 meter and radius of base 20 cm if it rotates at constant angular velocity 10 rev/sec about its principal axis? What angular velocity would be necessary to give it the same kinetic energy if it were rotated about an axis through the vertex perpendicular to the principal axis?

50. Two particles A and B with masses m_A and m_B respectively move subject only to their mutual interaction. What is true of their instantaneous velocities? Specialize to the case where they are initially at rest.

If A moves in a circle of radius r_A with constant angular velocity, show that B must likewise move in a circle with the same angular velocity. What is the radius of this second circle? Specialize to the lowest energy state in the Bohr model of the hydrogen atom and give numerical values.

51. Two particles B and C with masses m_B and m_C respectively move subject only to their mutual interaction. Find the expression for the acceleration of C relative to B in terms of the acceleration of C in the reference frame. Find the relative fractional acceleration of C. What are its maximum and minimum possible values?

What will be the effective relative mass., i.e., the mass of a particle which, moving with the *relative* acceleration, has the same unbalanced force acting on it as acts on B and C themselves?

- 52. Prove that $\int_{V_0}^{V_1} p \, dV$, where p is the excess pressure and V the volume, represents the work required to compress a fluid from volume V_0 to volume V_1 . If the fluid is an ideal gas and the compression is adiabatic, calculate the work. If the fluid is a liquid, obtain an expression for the work by the use of the bulk modulus.
- 53. The energy of a single molecule of a fluid when in a state of compression from volume V_0 to volume V may be expressed in terms of the series

$$E = E_0 + (V - V_0) \left(\frac{dE}{dV}\right)_0 + \frac{(V - V_0)^2}{2} \left(\frac{d^2E}{dV^2}\right)_0 + \cdots,$$

where E_0 is the equilibrium energy in the uncompressed state and $\left(\frac{dE}{dV}\right)_0^{\circ}$ is the derivative referring to equilibrium conditions. If equilibrium corresponds to minimum molecular energy, what value must be assigned to the second term on the right? Derive the following expression for the bulk modulus of the fluid

$$k = V_0 N \left(\frac{d^2 E}{dV^2}\right)_0,$$

where N is the number of molecules in volume V_0 . Discuss the physical significance of the requirement which the positive nature of k places on $\left(\frac{d^2E}{dV^2}\right)_0$.

54. If in Problem 52 the change in volume from V_0 to V is that involved in the change from liquid to solid state, the work done may be expected to be approximately equal to the heat of fusion. Test this hypothesis by application to the following substances: lead, bismuth, ethyl alcohol, and tin.

55. The critical temperature of a substance may plausibly be defined from the kinetic theory point of view as the temperature for which the average energy per molecule equals the energy required to separate the molecules completely. The latter may be calculated from the result of Problems 52 and 53. Thus get an estimate of the order of magnitude of the critical temperature of water on this hypothesis.

56. A possible equation of state of a liquid has been given in the form

$$p(V - V_0) = 3NkT,$$

where V_0 is the *minimum* possible volume for the closest packing of the molecules. The other symbols have their usual significance as in Chapter VI. How does the pressure vary with depth for such a liquid under isothermal conditions?

57. From the standpoint of the quantum theory light is looked upon as being corpuscular in nature and composed of particles called "photons," which have no mass but carry energy h_{ν} and momentum h_{ν}/c , where ν is the frequency of the light, c is the velocity of light and h is Planck's constant (see Sec. 12.4). Thus any change in the energy and momentum of a photon involves an alteration in the frequency of the corresponding light. Suppose that a photon of frequency ν_0 moving along the positive x axis collides with an electron at rest with respect to a certain set of rectangular axes. After collision the electron moves off with velocity v in a direction making angle θ with the positive x axis, while the photon travels in a direction making angle ϕ with the same axis. The mass of the electron at rest is m_0 , while the mass in motion with velocity v is $m = m_0/\sqrt{1-v^2/c^2}$ (from the theory of relativity. See Sec. 12-4). The kinetic energy of the electron, which in this case is its total energy, on the theory of relativity is $(m - m_0)c^2$. Set up the equation which expresses the conservation of energy for the system consisting of the photon and electron, and the two equations expressing the conservation of momentum of the system along and perpendicular to the x axis. Solve the three equations for ν , In particular find v, and θ in terms of v_0 , m_0 , and ϕ .

$$\nu = \frac{\nu_0}{1 + h\nu_0/m_0c^2 \cdot (1 - \cos\phi)}.$$

Hence find for the change in wave length of the photon $(\lambda = c/\nu)$ due to the collision:

$$\Delta\lambda = \frac{c}{\nu} - \frac{c}{\nu_0} = \frac{h}{m_0 c} (1 - \cos \phi).$$

The so-called Compton effect in the scattering of X-rays by light elements is described by the theory here indicated (see Ruark and Urey, "Atoms, Mole-

cules and Quanta," p. 84 ff.). It is interesting to observe that in this collision problem we invoke the conservation of energy as well as of momentum, in contrast to the assumption used in the study of collision problems in classical mechanics.

- 58. Discuss the motion of the bob of a simple pendulum in a medium which resists the motion with a force varying directly as the velocity. Also discuss the case where the resisting force varies as the square of the velocity.
- 59. A bead moves on a smooth circular wire of radius a; the plane of the wire is inclined at an angle ϕ to the vertical. Find the frequency of the small oscillations of the bead about the lowest point of the wire.
- 60. The motion of the bob of a simple pendulum when first observed has an amplitude of 6 cm. Eight minutes thereafter the amplitude of swing has decreased to 4 cm. Find the decay modulus. How long a time must elapse before the amplitude has diminished to 2 mm? If the length of the pendulum is 50 cm, calculate its natural frequency and compute the error involved in neglecting the effect of the dissipation. What is the logarithmic decrement?
- 61. In a special arrangement of Atwood's machine both particles have the same mass, so that the system is originally in equilibrium. One of the particles is then pulled aside a short distance in the original plane of the system and allowed to swing in this plane as a simple pendulum. Discuss its motion.
- 62. A perfectly flexible string of length l and line density ρ is stretched with tension τ . Show that when it is displaced transversely, the total kinetic energy and total potential energy (neglecting dissipation) may be expressed in the following way

$$T = \frac{1}{2} \int_0^l \rho \dot{\xi}^2 dx,$$

$$V = \frac{1}{2} \int_0^l \tau \left(\frac{\partial \xi}{\partial x} \right)^2 dx.$$

Here ξ represents the displacement at distance x along the string from one end chosen as origin. Taking the solution for a progressive wave along the string, i.e., $\xi = A \cos(\omega t - \kappa x)$ (from Sec. 10.5), evaluate T and V.

63. Consider a perfectly flexible circular membrane of radius a which is fixed at the periphery. Imagine it depressed by a uniform pressure p. If the superficial tension, i.e., force per unit length (Sec. 11.8), is τ (assumed to remain constant under the displacement), show that the displacement of a point on the membrane distant r from the center is given approximately by

$$\xi = \frac{pa^2}{4\tau} \left(1 - \frac{r^2}{a^2} \right).$$

Hint: Discuss the static equilibrium of the ring element $2\pi r dr$. What is the maximum displacement ξ_0 and what is the shape of the deformed membrane?

64. Find the expression for the total kinetic energy of the membrane in the previous problem (Problem 63) if it is depressed and let go, assuming that ξ is always given by the expression above, and that the instantaneous velocity of the center is $\dot{\xi}_0$. In this way prove that the *effective* mass of the membrane is *one-third* of the actual mass. N.B. The effective mass, m_e , is that for which $\frac{1}{2}m_e\dot{\xi}_0^2$ is equal to the total kinetic energy.

65. Referring again to Problem 63 show that the potential energy of the circular membrane to the approximation there assumed is

$$V=\frac{8\pi\tau}{a^2}\left(\frac{a^2}{4}\right)\frac{\xi_0^2}{2},$$

so that the effective stiffness coefficient of the membrane is $f=2\pi\tau$. (N.B. The effective stiffness coefficient is that such that $\frac{1}{2}f\xi_0^2$ is equal to the total potential energy.) Hence show that the natural frequency of vibration of the

membrane is $\nu_0 = \frac{1}{2\pi} \sqrt{\frac{6\tau}{\rho a^2}}$, where $\rho =$ density of membrane material. It

should be noted that the last three problems refer to an approximate theory of the circular membrane. According to the more exact theory, the membrane has a whole set of natural frequencies (recall the stretched string) of which ν_0 is a first approximation to the fundamental. (See Stewart and Lindsay, Acoustics, p. 202.)

66. A perfectly flexible string of length l is stretched in a horizontal line with tension τ and loaded at equal intervals a with n particles of mass m, so that (n+1)a=l. One of the masses is displaced slightly in the vertical plane. Using the general method of Sec. 9-8 write the equations of motion of the system, assuming that the mass of the string itself is negligible. Show that if the vertical displacements of the particles are denoted by $q_1 \dots q_n$ respectively,

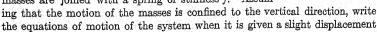
$$q_j = \sum_{k=1}^n A_k \sin \frac{(j-1)k\pi}{n+1} \cdot \cos (\omega_k t + \epsilon_k),$$

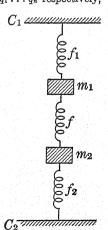
where j runs from 1 to n, satisfies the equations, provided that

$$\omega_k = 2\sqrt{\frac{\tau}{ma}}\sin\frac{k\pi}{2(n+1)}.$$

The latter are known as the *characteristic frequencies* of the loaded string. Compare with the special case of Sec. 9.8. Take a special case and calculate numerical values.

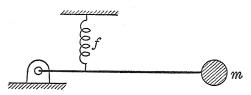
67. In the diagram m_1 and m_2 are masses joined to the rigid supports C_1 and C_2 respectively by the two springs with stiffness f_1 and f_2 respectively. Moreover, the masses are joined with a spring of stiffness f. Assum-





from equilibrium. Neglect the mass of the springs. Determine the frequencies of the motion.

68. The diagram indicates a massless, inflexible rod, 1 meter in length, hinged at one end and carrying a mass m=500 grams at the other end. The rod is supported by a vertical spring of stiffness 1 kilogram wt per cm attached



25 cm from the hinge. When the mass m is displaced vertically a *small* distance and let go, the amplitude of motion is found to be reduced to 1/eth of its original value in 10 sec.

- (a) What is the natural frequency of the vibrations of m?
- (b) What is the logarithmic decrement of the motion of m?
- (c) If the system were to be driven at its resonance frequency by a periodic force with a maximum value of 100 grams wt, at what rate would it dissipate energy?
- 69. Consider a pair of coupled oscillators like those described in Sec. 9.8 and for which the displacements are given by

$$\xi_1 = A \cos (\omega_1 t + \epsilon_1) + B \cos (\omega_2 t + \epsilon_2),$$

$$\xi_2 = A \cos (\omega_1 t + \epsilon_1) - B \cos (\omega_2 t + \epsilon_2).$$

Suppose the system is started from rest by giving the first component the initial displacement ξ_{10} while the second has zero displacement. Show that the resultant motion can be expressed in the form

$$\xi_1 = \xi_{10} \cos \frac{(\omega_1 + \omega_2)}{2} t \cdot \cos \frac{(\omega_1 - \omega_2)}{2} t,$$

$$\xi_2 = -\xi_{10} \sin \frac{(\omega_1 + \omega_2)}{2} t \cdot \sin \frac{(\omega_1 - \omega_2)}{2} t.$$

Discuss the physical significance of this result.

- 70. The point of suspension of a simple pendulum moves with constant velocity in a circle in the plane of the pendulum's motion. Write the equations of motion and solve them for the case in which the circle is very small.
- 71. A condenser of capacitance C is charged through a resistance R by a constant source of e.m.f. E. Find the relaxation time, i.e., the time required for the charge on the condenser to rise to within 1/e of the static value.
- 72. Write the relaxation equation (Cf. eq. 9.2–47) corresponding to the imposition of a constant excess pressure p_e on a fluid of mean density ρ_0 . This

is the differential equation for the excess density ρ_e . Find the expression for the relaxation time.

Do the same problem when the excess pressure is a harmonic function of the time and solve to find p_e/ρ_e in terms of the relaxation time. What physical significance can be associated with this ratio? (Cf. eq. 11.6-37.)

73. A uniform steel cylindrical rod of length l_1 and radius a is suspended horizontally at its mid-point from a steel wire of length l_2 and radius b. Find the ratio of the number of oscillations per second of the bar (eq. 10-2-18) at temperature T_1 to the number at temperature T_2 . Consult physical tables for necessary numerical data.

74. Given two homogeneous media separated by a plane boundary. Let us suppose that all particles move in the first medium in straight line paths with constant velocity V_1 and in the second medium with constant velocity V_2 . Find the path of a particle between point A of the first medium and point B of the second medium satisfying the condition that it shall be traversed in the minimum time. (Note: This is the famous principle of Fermat for the propagation of light. It should be shown that Snell's law of refraction follows from the condition cited.)

75. A cubical box of side a is supposed to be filled with electromagnetic radiation in the form of plane harmonic waves which, however, must satisfy the boundary condition that the displacement is zero on all faces of the box. Show that the allowed frequencies are given by

$$\nu = \frac{c}{2a} \sqrt{n_1^2 + n_2^2 + n_3^2},$$

where c = velocity of the waves, and n_1 , n_2 , n_3 form a set of any three integers. Hint: Use the expression for the characteristic frequencies of a stretched string (Sec. 10.5).

76. Referring to the previous problem (Problem 75) prove that the number of possible modes of oscillation inside the cubical space having frequencies included in the frequency region ν to $\nu + \Delta \nu$ is given by $4\pi a^3/c^3 \cdot \nu^2 \Delta \nu$. Hint: Represent each possible frequency by a point in a three-dimensional lattice space with coördinates the integers n_1 , n_2 , n_3 . The number of such points lying in any given volume of this space will be numerically equal to the volume, approximately. If each mode of oscillation has attached to it the amount of energy kT (where T is the absolute temperature and k is Boltzmann's gas constant), show that the energy density of the radiation in the frequency inter-

val ν to $\nu + \Delta \nu$ is $\Delta E_{\nu} = \frac{8\pi \nu^2 kT}{c^3} \Delta \nu$. Express this in terms of wave length.

This gives the Rayleigh-Jeans law for the frequency distribution of energy in the radiation from a perfect radiator of electromagnetic radiation (i.e., a black body). Unfortunately the result does not agree with experiment. This discrepancy was instrumental in the introduction of the quantum theory by Planck in 1900.

77. The displacement in a plane harmonic acoustic wave progressing in the x direction may be written in complex form as $\xi = Ae^{i(\omega t - \kappa x)}$ where A is the amplitude, $\omega = 2\pi\nu$, where ν is the frequency and $\kappa = \omega/c$, where ν is the wave velocity. Find the expression for the ratio ν/x , where ν is the excess pressure in the wave and ν/x (the so-called volume current) = ν/x , ν/x being the area of cross-section of the wave front. This ratio is called the acoustic impedance of the wave. Hint: Use the definitions and theory in Sec. 11-6.

78. A plane harmonic acoustic wave $\xi = Ae^{i(\omega t - \kappa x)}$ progresses through a cylindrical tube of cross-sectional area S_1 in the x direction (i.e., from left to right). At a certain point there is an abrupt change in the area of cross-section from S_1 to S_2 . Show that the fraction of the incident energy at the boundary

which is reflected is $\left(\frac{S_2 - S_1}{S_2 + S_1}\right)^2$. For this purpose utilize the boundary con-

ditions that at the boundary there exists continuity in pressure p as well as in volume current X. Assume that there is no wave in the negative x direction in the tube to the right of the boundary.

79. Show that the wave equation for spherical acoustical waves diverging from or converging to a point may be written in the form

$$(\dot{r}\phi) = \dot{c^2} \frac{\partial^2(r\phi)}{\partial r^2},$$

where ϕ is the velocity potential and r the distance from the point in question. Hence show that the general solution has the form

$$\phi = \frac{1}{r} [f_1(ct - r) + f_2(ct + r)].$$

Considering the velocity potential for spherical harmonic waves in the form

$$\phi = \frac{A}{r} e^{i(\omega t - \kappa r)},$$

deduce the expression for p/\dot{X} (see problem 77), i.e., the acoustic impedance, and compare with that for a plane wave. What is the limiting value of the acoustic impedance of the spherical wave as $r \to \infty$?

- 80. A torsional wave travels along an infinitely long uniform cylindrical solid rod. Show that the velocity of propagation is $\sqrt{\mu/\rho}$, where μ is the shear modulus and ρ is the density.
- 81. A conical horn is in the form of the frustum of a right circular cone with slant height h. If both ends of the horn are open, show that its characteristic or resonance frequencies for spherical acoustic waves are given by v = nc/2h where n is any integer. Compare with the case of the stretched transversely vibrating string and with the cylindrical organ pipe.
- 82. Harmonic acoustic waves pass through a tube with rigid walls and of varying cross-section S (i.e., a horn). The diameter is supposed to be everywhere smaller than the wave length. Show by setting up the equation of motion and the equation of continuity for the fluid (e.g., air) in the tube and

by making the usual acoustic approximations (Sec. 11-6) that the approximately valid differential equation for the acoustic excess pressure in the tube is

$$\frac{\partial^2 p}{\partial x^2} + \frac{1}{S} \frac{dS}{dx} \frac{\partial p}{\partial x} + \kappa^2 p = 0,$$

where $\kappa=2\pi/\lambda=2\pi\nu/c$ in the usual notation. Solve this equation for the special case of a conical tube. This differential equation is basic for much of modern acoustical horn theory.

 γ 83. A semi-infinite solid rod of cross-sectional area S is traversed by harmonic compressional elastic waves of frequency ν in both directions. Show that the complex volume current X (i.e., product of particle velocity ξ and area S) and the complex excess stress T (due to the waves) at any point may be represented by

$$\dot{X} = \left[\dot{X}_1 \cos \kappa x + \frac{iT_1S}{\rho_0 c} \sin \kappa x \right] e^{i\omega t},$$

$$T = \left[T_1 \cos \kappa x + \frac{i\dot{X}_1\rho_0 c}{S} \sin \kappa x \right] e^{i\omega t},$$

where \dot{X}_1 and T_1 are the complex volume current and excess stress at x=0 respectively, $\rho_0=$ mean density of the material of the rod and c= velocity of the waves. As usual, $\kappa=2\pi/\lambda=\omega/c=2\pi\nu/c$.

84. The solid rod in the preceding problem (Problem 83) is loaded at equal intervals of length 2l with concentrated weights of mass m. By setting up the usual boundary conditions (see Problem 78 above) involving continuity of pressure and volume current, and making use of the recurrent nature of the structure show that, if X_n and T_n are the volume current and excess stress halfway between the (n-1)st and nth loads, we have the relations

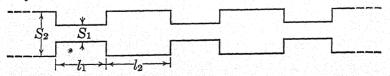
$$\dot{X}_{n+1} = \dot{X}_n e^{-iW},$$

$$T_{n+1} = T_n e^{-iW},$$

where $\cos W = \cos 2\kappa l - \frac{m\omega}{2\rho_0 cS} \sin 2\kappa l$. Hence show that the structure behaves like a low frequency pass filter for compressional waves, allowing the free transmission only of those waves whose frequencies satisfy the condition

$$-1 \le \cos W \le 1$$
.

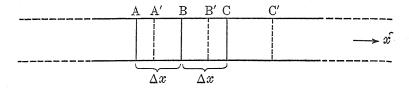
85. The accompanying schematic diagram represents an infinite succession of cylindrical tubes of cross-sectional area S_1 and S_2 respectively and lengths



 l_1 and l_2 respectively. If plane harmonic waves travel along this structure in both directions, show by the use of boundary conditions as in Problem 78 that

the structure is a low pass filter and find the conditions for transmission and attenuation bands.

86. The accompanying figure represents a section of an elastic fluid medium extending in the x direction. Consider the two elements of length AB and BC both equal to Δx . If the medium is given a small displacement so that the



layer at A moves through ξ , show that there is a change in density such that, in the displaced element A'B', the density is to a first approximation

$$\rho = \frac{\rho_0}{1 + \frac{\partial \xi}{\partial x}},$$

where ρ_0 is the mean equilibrium density. Find $d\rho$ and (by the use of Hooke's law for longitudinal extension) deduce the equation of motion in the x direction in the form

$$\ddot{\xi} = Y/\rho_0 \cdot \frac{\partial^2 \xi}{\partial x^2} / \left(1 + \frac{\partial \xi}{\partial x}\right)^2 \cdot$$

To what form does this reduce when $\left|\frac{\partial \xi}{\partial x}\right| \ll 1$? Discuss the physical significance of this.

87. Two waves of intensity I_1 and I_0 respectively are said to differ by D decibels if

$$D = 10 \log_{10} I_1 / I_0.$$

When the intensity of a wave diminishes through absorption by the medium, the intensity I at distance r from the place where the intensity is I_0 is given by

$$I = I_0 e^{-\alpha r}$$

Find the loss in decibels per cm (db/cm) in terms of α . Apply to Problem 29 of Chapter XI.

^{88.} The surface of a liquid is in the form of a cylinder of length h and radius R. Show that the internal pressure necessary to keep the liquid cylinder in equilibrium is T_s/R , where T_s is the surface tension.

89. An air bubble is formed in water by forcing air through a vertical tube whose open end is 2.91 cm below the water surface. The radius of the tube orifice is 0.025 cm. If the excess pressure in the bubble at the moment of its breaking away, as measured by a manometer, is 9.15 cm. of water, find the surface tension of the water.

90. A thin circular disk of radius a is suspended so that its lower surface touches a liquid which wets it. Show that the force with which it is necessary to pull upward to remove the disk from the liquid is given by $2\pi a^2 \sqrt{\rho g T_s}$, where ρ is the density of the liquid, g is the acceleration of gravity and T_s is the surface tension. Carry out the numerical calculation in the case of water for a disk of radius 10 cm.

91. According to Rayleigh ("Theory of Sound," par. 353) the velocity of surface waves of moderate amplitude in an ideal, incompressible fluid under gravity and surface tension is given by

$$c = \sqrt{(g/k + T_s k/\rho) \tanh kl},$$

where l= depth of fluid, $T_s=$ surface tension, $\rho=$ fluid density, $k=2\pi/\lambda$, where $\lambda=$ wave length, and g= acceleration of gravity.

If the fluid is water, find the expression for the group velocity for the following special cases: $kl \ll 1$; and $kl \gg 1$. Specialize the second case to both large and small λ . Introduce numerical values where possible. Comment on the physical significance of the results.

92. Show that the differential equations of the stream lines of a fluid in steady flow are

$$\frac{dx}{u} = \frac{dy}{v} = \frac{dz}{v}$$

where u, v, w are the rectangular components of flow velocity. If one confines one's attention to two-dimensional flow in the xy plane, find the stream lines corresponding to u = Cy and v = -Cx, where C is a constant.

93. How will the Poiseuille formula for steady flow of a viscous fluid through a cylindrical tube (eq. 11.7-8) be modified if the viscosity is a linear function of the flow velocity, i.e., $\eta = \eta_0 + Cu$, where C is a constant?

94. Starting with the Newtonian equation of motion for the case of one degree of freedom and assuming that the force is conservative, derive Hamilton's principle, i.e., show that

$$\delta \int_{t_0}^{t_1} \left(T - V\right) dt = 0,$$

the variation having the meaning of Sec. 12.1.

95. Show that Hamilton's principle for a conservative system may be stated in the following form: A conservative dynamical system moves in such a way that the time average of the potential energy over any time interval differs least (or most) from the time average of the kinetic energy over the same interval. Find the expressions for $\overline{T} - \overline{V}$ in the special cases of the harmonic oscillator, motion in a uniform field and motion in a central inverse square field.

96. The Hamiltonian for a simple harmonic oscillator of mass m and stiffness k is (see Sec. 12·2 and Problems 8 and 9 at the end of Chapter XII)

$$H = \frac{1}{2m} p^2 + \frac{k}{2} q^2,$$

where q is the coördinate and p the conjugate momentum. Introduce the transformation from the coördinates p, q to P, Q where

$$q = \sqrt{\frac{2}{m\omega}} \sqrt{P} \sin Q,$$

$$p = \sqrt{2m\omega} \sqrt{P} \cos Q.$$

Here ω is a parameter. Find the form of the Hamiltonian in the new coordinates, and show that the canonical equations of motion

$$\frac{\partial H}{\partial q} = -p, \, \frac{\partial H}{\partial p} = \dot{q}$$

now become

$$P = \alpha$$
, $Q = \omega t + \beta$,

where α and β are arbitrary constants and $\omega = \sqrt{k/m}$. Hence show that the solution to the problem is

$$q = \sqrt{\frac{2\alpha}{m\omega}} \sin \left(\sqrt{\frac{k}{m}} \, t + \beta \right) \cdot$$

The transformation here considered is called a *canonical transformation* and is of much importance in problems of atomic mechanics.

97. From the theory of relativity show that the total energy E of a free particle can be written in the form

$$E^2 = p^2c^2 + m_0^2c^4,$$

where p= the momentum mv, and m is the variable mass. What is the momentum of an electron with 2×10^6 electron volts energy? (The electron volt is the energy gained by an electron in falling through a potential difference of 1 volt.)

98. Write and solve Lagrange's equations for a particle which moves in a plane subject only to a constant force in the x direction.

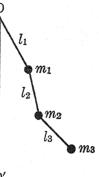
99? Mass particles m_1 , m_2 , m_3 are attached to a vertical flexible string as indicated in the diagram. If the particles are displaced arbitrarily from the

vertical equilibrium position (in the same plane) and released, write Lagrange's equations for the subsequent motion.

100. Show from the discussion in Sec. 12.4 that the wave mechanical Schrödinger equation for a particle moving along the x axis in a constant potential field (i.e., no resultant force) is of the form

$$\frac{d^2\psi}{dx^2} + \frac{8\pi^2 m}{h^2} (E - V)\psi = 0,$$

where V = constant potential energy and E = total energy. Solve this equation and find the values of E which are allowed by the condition that $\psi = 0$ at x = 0 and x = l.





ANSWERS TO SELECTED PROBLEMS

CHAPTER I

1. (a)
$$\sqrt{(x_2-x_1)^2+(y_2-y_1)^2}$$
; (b) $\sqrt{r_1^2+r_2^2-2r_1r_2\cos(\theta_1-\theta_2)}$; (c) $\sqrt{(x_2-x_1)^2+(y_2-y_1)^2+2(x_2-x_1)(y_2-y_1)\cos 60^\circ}$.
2. $x'=x\cos\theta+y\sin\theta$; $y'=y\cos\theta-x\sin\theta$.

5. $4\pi^2\left(1+\frac{1}{8\pi^2}\right)$ meters/sec², approximately.

7. Angle with the north is 27°41'.

10. 50 cm/sec; $100/\sqrt{2}$ cm/sec.

12. $4.1 \times 10^{16} \, \text{rad/sec.}$

15. Velocity after collision = $\frac{100\sqrt{10}}{3}$ cm/sec in direction making angle $\theta = \arctan \frac{1}{3}$ with y axis. Initial kinetic energy $= \frac{11}{4} \times 10^6$ ergs. Final kinetic energy = $\frac{5}{3} \times 10^6$ ergs.

16. Acceleration of m_1 relative to $m_2 = a_1(1 + m_1/m_2)$, where a_1 is acceleration of m_1 in the fixed reference system.

CHAPTER II

- 3. $\frac{m}{k}\sqrt{2gh}$; 4.605m/k.
- 6. If particle starts from rest at the origin,

$$x = \sqrt{F/k} \log \cosh t;$$
 $v = \sqrt{F/k} \tanh t,$

if k =resisting force coefficient.

- 8. Simple harmonic; $\pi \sqrt{3/4\pi G\rho}$; where ρ is the density of the sphere and G the constant of gravitation.
- 10. $\frac{1}{2\pi}\sqrt{g/h}$.
- 12. 4.8 $\sqrt{N} \times 10^9$ cm/sec.
- 15. Time is $h^2 \sqrt{m/k}$.

CHAPTER III

1. 78 sec; 1.7×10^5 ft.

Total length of path

$$= \frac{-v_0^2 \cos^2 \theta}{2a} \left[u \sqrt{1 + u^2} + \log \left(u + \sqrt{1 + u^2} \right) \right]_{\tan \theta}^{\tan \theta - gR/\tau_0^2 \cos^2 \theta}$$

3. $v_0/2g \sin \alpha$.

- 5. Mass of sun: 2×10^{33} grams (approximately); mass of earth: 6×10^{27} grams (approximately).
- 6. Ellipse with center at force center.
- 7. Inverse cube.

mass M.

430

- 10. 2×10^{80} grams.
- 15. Inverse 5th power attraction.

CHAPTER IV

- 1. $\frac{1}{2}(M + \rho l/3)\xi_M^2$, where ξ_M is the velocity of the particle.
- 3. $x_{\text{max}} = 20 \text{ cm}$; $\dot{x}_{\text{max}} = 200 \sqrt{10} \text{ cm/sec.}$ Curve is an ellipse with center at origin.
- 5. $K_{\rm aph} = \frac{1}{2} \frac{mk^2}{r_{\rm max}^2}$; $K_{\rm peri} = \frac{1}{2} \frac{mk^2}{r_{\rm min}^2}$, where $k = {\rm area}$ constant $= ac(1 \epsilon^2)/m$, where $a = {\rm semi-major}$ axis and $\epsilon = {\rm eccentricity.}$ $c = Ne^2$ for electron moving about nucleus with charge Ne, or GmM for gravitational attraction to
- 6. 54 electron volts (it is assumed that the atom is in its lowest energy state to begin with).
 Energy states corresponding to principal quantum numbers 6 and 4 respectively.
- 7. $1.3m \times 10^{11}$ ergs; $3.1m \times 10^{11}$ ergs. Potential at $P = -5.1 \times 10^{11}$ ergs. Potential at center of earth $= -9.5 \times 10^{11}$ ergs.
- 9. (b) $V_0 e^{-\alpha r^2} (1 2\alpha r^2)$; (c) It will not reach the origin; (d) zero; (e) simple harmonic motion.
- 10. $2GM \log r/a + V_a$, where V_a is the potential at the surface.
- 12. $(a^2/4 + r^2 + ar \cos \theta)^{-1/2} (a^2/4 + r^2 ar \cos \theta)^{-1/2} = C/e$ where C is a constant, and polar coordinates r and θ are employed with the origin taken mid-way between the charges. These are the traces of the family on any plane through the line joining the charges. The surfaces are obtained by revolving these traces about the line.
- 16. $V = MGe^{-\alpha r} \left[\frac{\alpha}{2} + \frac{1}{r} \right] \frac{GM}{r}$; $M = \text{total mass} = 8\pi \rho_0 / \alpha^3$.

CHAPTER V

- 1. Infinite.
- 3. $mg \sin \theta$; $mg \cos \theta$.
- 5. Approximately 0.9.
- 7. $\cos \theta_1 \cos \theta_2 = C = \text{constant}$. The angle θ_1 is the angle which the line from any point on the line of force to the north pole of the magnet makes with the magnet. Similarly for θ_2 .
- 9. Approximately 2.52 cm.

12. If α_1 is the angle which Dm_1 makes with CA and α_2 the angle which Dm_2 makes with CB, equilibrium holds for the values of α_1 and α_2 satisfying the relations

$$m_1 \sin \theta_1/\cos \alpha_1 = m_2 \sin \theta_2/\cos \alpha_2$$
,
 $\sin \phi_1/\sin \alpha_1 + \sin \phi_2/\sin \alpha_2 = l/h$,

where
$$l = \text{total length of string and } \phi_1 = \angle DCA, \phi_2 = \angle DCB.$$

.14. $\rho = \frac{C}{g\sqrt{a^2 + x^2}}$, where C = constant horizontal tension and a is parameter of parabola, i.e., $y = x^2/2a$ is equation of parabola.

17. Approximately 1 ft.

CHAPTER VI

2. The distance of the center of mass of the system at time t is given by $\bar{x} = (v_1 - v_2)t/2 + (x_1 + x_2)/2$. The particles will come to rest when their separation is

$$\frac{8e^2}{m(v_2+v_1)^2/2+8e^2/(x_2-x_1)}.$$

They will thereafter separate.

4. Distance of the center of mass from the pulley at time t is

$$\frac{m_1x_{10} + m_2x_{20}}{m_1 + m_2} + \frac{1}{2} \left(\frac{m_2 - m_1}{m_2 + m_1}\right)^2 gt^2,$$

where x_{10} and x_{20} are initial distances from the pulley of m_1 and m_2 respectively.

Kinetic energy of
$$m_1 = \frac{1}{2}m_1 \left(\frac{m_2 - m_1}{m_2 + m_1}\right)^2 g^2 t^2$$
;

Kinetic energy of
$$m_2 = \frac{1}{2}m_2 \left(\frac{m_2 - m_1}{m_2 + m_1}\right)^2 g^2 t^2$$
;

Potential energy of
$$m_1 = m_1 g \left[h + \frac{m_2 - m_1}{m_1 + m_2} \cdot \frac{gt^2}{2} \right];$$

Potential energy of
$$m_2 = m_2 g \left[h - \frac{m_2 - m_1}{m_1 + m_2} \cdot \frac{g t^2}{2} \right]$$
;

It is here assumed that $m_2 > m_1$.

- 5. Parabola with projection speed $\sqrt{4v^2 + v_1^2}$ and projection angle with horizontal arc tan $v_1/2v$, where v = speed of shell at moment of explosion and $v_1 =$ initial downward speed of piece falling vertically.
- 8. $4\pi^2a^3 \not G(m_1 + m_2)$, where a is the semi-major axis of the elliptical orbit.
- 10. Time is very closely $2\sqrt{100/g} \cdot \log_e 10$. Velocity is $\sqrt{100g}$.

12.
$$\cos \theta = \frac{n^2 - n_{\theta}^2 - n_{s}^2}{2n_{\theta}n_{s}}$$
; $\cos \theta = \pm 1$.

15. 21.3 cm.

18.
$$U_n = \frac{-2\pi^2 M e^4}{n^2 h^2}$$
, where $M = \text{reduced mass} = \frac{m_e m_n}{m_e + m_n}$, with $m_e = \text{mass of electron}$, $m_n = \text{mass of nucleus}$.

CHAPTER VII

- 1. $\mathbf{v}_P = \mathbf{v} + \boldsymbol{\omega} \times \mathbf{r}$; circular helix.
- 2. 1.6×10^9 ergs.
- 3. Frequency = 0.7 vib/sec; torque = $2.1 \times 10^6 \text{ cm dyne}$.
- 4. 12.6 rad/sec.
- 5. $E_n = n^2 h^2 / 4\pi^2 m l^2$.
- 7. $\frac{5}{7}g\sin\theta$, if θ is the angle of inclination of the plane.
- 17. On axis of symmetry at distance $4\sqrt{2r/9\pi}$ from center of disc.
- 34. $x^2/a^2 + y^2/b^2 + z^2(1/a^2 + 1/b^2) = C$, where C is constant.
- 20. On axis of symmetry of octant $3\sqrt{3}R/8$ from the center of the sphere.
- 29. 0.1 ft.
- 32. $x^2 + y^2 = A^2 \cos^2 \omega t$, if the motion takes place in the xy plane, and x, y are coordinates with reference to axes fixed in turn table. The frequency of the pendulum is $\omega/2\pi$, and the amplitude of its motion is A.

CHAPTER VIII

- 1. $1.845 \sqrt{R/g}$.
- 4. Time for descent = $\pi \sqrt{a/g}$.
- 5. If θ is angle which radius vector to point where particle leaves circle, makes with the vertical, then

$$\cos \theta = \frac{2}{3}$$
.

10.
$$a_1 = \text{acceleration of } m_1 = \frac{(m_1 gR - m_2 gr)R}{R^2(m_1 + M/2) + r^2(m_2 + m/2)}$$
.

(N.B. This assumes that the wheel and axle are homogeneous solid discs.)

CHAPTER IX

- 4. 14 cycles/sec; 3.92×10^5 dynes/cm; 20 dyne sec/cm; 1.4×10^{-2} .
- 1. $\frac{s}{K} = fA^2/3$, $\frac{s}{V} = fA^2/6$, where f = stiffness.

3.
$$\nu = \frac{1}{2\pi} \sqrt{\frac{f_1 + f_2}{m}}$$
; $\Delta \nu / \nu = \frac{1}{2} \left(\frac{\Delta f_1 + \Delta f_2}{f_1 + f_2} \right)$.

- 6. Decay modulus = 23 sec; log-decrement = 1.7×10^{-4} ; $\omega/\omega_0 = 1 \pm 2.5 \times 10^{-5}$. For air resonator $\omega/\omega_0 = 1 \pm 4.8 \times 10^{-5}$.
- 7. 3.26 cm.

9.
$$\nu_{\rm res} = 400/2\pi$$
 cycles/sec;

$$\nu_{\text{nat}} = \frac{1}{2\pi} (400 - 12.5 \times 10^{-4}) \text{ cycles/sec.}$$

13.
$$\nu = \sqrt{\frac{1}{2}(\nu_1^2 + \nu_2^2) \pm \sqrt{\frac{1}{2}(\nu_1^2 - \nu_2^2)^2 + 4\mu^4}}$$

where

$$\begin{split} \nu_1 &= \frac{1}{2\pi} \sqrt{\frac{C+C_1}{CC_1L_1}} \; ; \qquad \nu_2 = \frac{1}{2\pi} \sqrt{\frac{C+C_2}{CC_2L_2}} \; ; \\ \mu &= \frac{1}{2\pi} \sqrt{\frac{1}{C\sqrt{L_1L_2}}} \; . \end{split}$$

15. Amplitudes of m_1 and m_2 respectively in steady state are

$$A_1 = rac{1 - m_2 \omega^2 / f_2}{(1 - m_2 \omega^2 / f_2) (1 + f_2 / f_1 - \omega^2 m_1 / f_1) - f_2 / f_1};
onumber \ A_2 = rac{1}{(1 - m_2 \omega^2 / f_2) (1 + f_2 / f_1 - \omega^2 m_1 / f_1) - f_2 / f_1}.$$

16.
$$\frac{16}{3} \pi^4 \frac{e^2}{c^3} \nu^3 A^2$$
; number of periods = $3mc^3/16\pi^2 \nu e^2$.

CHAPTER X

2. 3×10^9 dynes.

4. Torque = 19.2×10^4 cm dynes. Frequency = 0.75 vib/sec.

6. $\frac{1}{2}X^2/Y$.

8. If $\delta = 0$ at $t = t_0$, we have

$$\delta = \frac{X_0}{A} \left(1 - e^{-\frac{A}{B}(t - t_0)} \right), t_0 < t < t_1;$$

$$\delta = \frac{X_0}{A} \left(1 - e^{-\frac{A}{B}(t_1 - t_0)} \right). \quad e^{-\frac{A}{B}(t - t_1)}, t > t_1.$$

15. Fundamental = 256 cycles/sec.

$$a_n = \frac{4}{n^2 \pi^2} \sin n\pi/2.$$

18. $\xi = \xi_0 e^{-Rx/2c} \cdot \cos(\omega t - kx)$, where $k = \omega/c$; $x_0 = 2c/R$.

19. (1) $\nu = nc/2l$, where $n = 1, 2, 3 \cdots$, $c = \sqrt{Y/\rho}$; where Y = Y oung's modulus;

(2) $\nu = (2n-1)c/2l$, where $n = 1, 2, 3 \cdots$;

(3) $\nu = nc/2l$, where $n = 1, 3, 5, 7 \cdots$ (or all odd integers).

CHAPTER XI

3. $-\frac{4}{3}\pi R^3 \rho g$, where $\rho =$ density of water; $\frac{\pi R^3}{3} \rho g$, upward thrust; $\frac{5}{3}\pi R^3 \rho g$, downward thrust.

5. \(\frac{2}{3}\) of density of water.

7. $p = p_0 e^{-\rho_0 gz/p_0}$, where z measures height above the earth's surface, and p_0 , ρ_0 denote pressure and density respectively at the surface.

9. $p = p_0 + \rho_0 g(z + kz^2)$, if $\rho = \rho_0 (1 + kz)$.

13. Velocity of efflux = $\sqrt{2gx/(1-S_B^2/S_A^2)}$. Rate of descent of free surface = $\sqrt{2gx/(S_A^2/S_B^2-1)}$.

15. 9.13×10^4 gallons/day.

17. $c_t = c_0 \sqrt{1 + t/273}$, where c_t is velocity at t° C and c_0 the velocity at 0° C; independent of the pressure.

18. $A_r = A_i \left(\frac{r_{12}-1}{r_{12}+1}\right)$; $A_t = 2A_i/(r_{12}+1)$, where A_i , A_r , A_t are incident, reflected and transmitted displacement amplitudes respectively and

 $r_{12} = \rho_{02}c_2/\rho_{01}c_1,$

where c_1 and c_2 are sound velocities in the two media respectively and ρ_{01} and ρ_{02} are the mean densities.

24. $\frac{\partial u}{\partial t} = -\frac{1}{\rho} \frac{\partial p}{\partial x}$.

sity.

26. $\phi = -Ac\cos(\omega t - kx)$;

 $s = \text{condensation} = -Ak \sin (\omega t - kx);$ $p_e = \text{excess pressure} = -\rho_0 c^2 Ak (\omega t - kx), \text{ where } \rho_0 = \text{mean den-}$

CHAPTER XII

4. Rectangular: $p_x = mx$, $p_y = m\dot{y}$, $p_z = m\dot{z}$. Spherical: $p_r = m\dot{r}$, $p_\theta = mr^2\dot{\theta}$, $p_\varphi = mr^2\sin^2\theta\dot{\phi}$. Cylindrical: $p_r = m\dot{r}$, $p_\theta = mr^2\dot{\theta}$, $p_z = m\dot{z}$.

5. $mA^2\omega/4$; $mA^2\omega/4 + \frac{15}{16}\pi m\omega\lambda^2A^2$.

6. The mass 4m rises with acceleration $\frac{7}{27}g$.

INDEX

Absolute value of a complex number, 263

Absorption of acoustic waves by the medium, 368

-, of light by oscillators, 287

Acceleration, definition, 12
—, angular, 14, 178, 179, 192

-, average, 12

-, centripetal, 14, 71, 222

-, component, 59, 60, 62, 64, 70

-, Coriolis, 222, 224, 238

-, instantaneous, 12

-, of a particle, 12

—, of gravity, determination of, 37, 190

-, relative to moving axes, 222

-, tangential, 14

-, uniform, along straight line, 33

-, vector, resolution of, 58

Acoustic resonator, 281

-, effective mass in, 281

—, kinetic reaction in, 282

-, radiation of sound energy by, 283

—, resonance frequency of, 283

—, stiffness of, 282

Acoustic waves in fluids, 360

Action and reaction, 19, 21

Action, mutual, of two particles in definition of mass, 15

Addition of vectors, 7

Additive property of masses, 16

Adiabatic gas law, 357, 365

Adiabatic invariant, 296

Adsorption, 382

Aggregate of material particles, 141, 175; see System of particles.

Alpha particles, 87

-, angular deviation of, 89

-, deflection by atomic nuclei, 87

-, velocity of approach, 89

Amplitude, of a complex number, 263

—, of simple harmonic motion, 39, 265

-, of simple harmonic motion, in plane, 64

-, of simple harmonic wave, 320

Amplitude angle, 240

Analytical conditions for equilibrium

of a particle, 120

Analytical expression for condition of equilibrium of rigid body, 213

Angle, solid, 109

Angle of contact, 374

Angle of repose, 131

Angular acceleration, 14, 178, 179, 192

Angular deviation of alpha particle,

Angular momentum, 179, 193, 216, 218, 225, 228, 230

Angular motion, 11

—, measure of angular displacement, 12

Angular velocity, 11, 221

—, average, 11

-, uniform, 40

-, vector representation, 12

Anomaly, eccentric, 81

-, mean, 82

Aphelion in elliptical motion, 78

Apparent velocity of relative motion,

220

Archimedes' Principle, 341

Area constant, 71

Areas, law of, 71, 396

Argument of complex number, 263

Atom of hydrogen, 82

—, Bohr model of, 82, 101, 148, 152, 171, 286

-, excited state of, 102

Atom of hydrogen, mass of, 83 Atomic oscillations, 286 f. Atomic structure, 289 -, Bohr-Rutherford theory, 82 f., 399 -, wave mechanics theory, 399 Atomic theory, simple harmonic oscillator in, 285 Atoms, 2 Attraction in inverse square field, 73 -, gravitational, 45, 145 Attractive forces, 46 Atwood machine, 36, 252, 254 Auxiliary circle, 81 Average angular velocity, 11 Average velocity of displacement, 9 Averages of kinetic and potential energy, time, 271, 322 -, space, 273, 323 Avogadro's hypothesis, 163, 170 Axes, coordinate, 4 —, moving, 216, 219, 225 Axis, energy of rotation about an, 180 -, major and minor of ellipse, 64, 67, 78, 81 —, rotation about a fixed, 176

Balance, 208 —, torsion, 49, 310 Ballistic pendulum, 415 Balmer series, 102 Bernoulli's principle of virtual work. 132 Bernoulli's theorem, 352, 355 ff. Bohr-Rutherford theory of atom, 399 Bohr-Sommerfeld quantum conditions, 83 Bohr's postulates, 82, 83 Bohr's theory of the hydrogen atom, see Atomic Structure Boltzmann gas constant, 163 Boundary conditions, 29 Boundary conditions in atomic structure, 84, 405 -, in particle dynamics, 34, 113 -, in wave motion, 328

Boundary value problem, 113, 405

Boyle's law, 161, 357, 365

Bulk modulus, 282, 305 Buoyancy, 342 —, center of, 344

Canonical equations of motion, 407 Capillarity, 370 ff. Capillary, depression, 373, 375 —, rise, 373, 374 —, tube, 374 Catenary, 44, 127 Cathode ray oscillograph, 68 Cathode rays, 68 Cavendish experiment, 50 Celestial mechanics, 82 Center, of buoyancy, 344 -, of force, 73

—, of gravity, 202, 207

-, of population, 207 -, of pressure, 202, 382

Center of mass, 141, 142, 147, 152, 207

---, of homogeneous bodies, 196 ---, of non-homogeneous bodies, 199

— —, of plane figures, 197

——, of solids, 201

---, of surfaces, 200 ---, rotational motion about, in rigid bodies, 195

---, translational motion of, in rigid bodies, 176, 195

Center,

-, of oscillation of a physical pendulum, 189

-, of suspension of a physical pendulum, 189

Central force field, 69 ff.

-, energy integral in, 98

-, energy relations in, 97

-, general theorems about motion in, 69 ff.

-, inverse square, 73 ff.

-, Lagrange's equations for, 394

—, potential in, 103 ff.

—, time equation in, 72, 76 Centripetal acceleration, 14, 71, 222

-, force, 71

Centreid, 197 Characteristic, frequencies, 292 -, values, 405 Circuit, see Electric circuit Circular motion, 11 -, uniform, 13 Coefficient, damping, 262 -, of compressibility, 305 Coefficient of friction, 130 $\stackrel{\sim}{-}$, determination of, 131 -, table of values, 130 Coefficient, of restitution, 144, 145, 156 -, of stiffness, 261 —, of viscosity, 167, 366 ----, determination of, for liquids, 368 Collisions, 96, 143 ff., 156 Comets, 80 Commutative law for vectors, 25 Complex displacement, 273 Complex numbers used in oscillation analysis, 263 Component, acceleration, 59 f., 62, 64, 70 -, forces, 59 -, method of, in solving equilibrium problems, 117, 214 -, rectangular, of a vector, 5 -, velocity, 9 f. Compressibility, 305 Compressional waves in a fluid, 358, 360 Compressive stress, 301 Concentration of a fluid, effect on surface tension, 382 Condensation, 362, 372 Conditions, boundary, see Boundary conditions —, of equilibrium of a particle, 116, 119

-, of rotational equilibrium of a

-, of translational equilibrium of a

rigid body, 203

rigid body, 203

Conic sections, 65, 73

-, eccentricity of, 74

Conservation of mechanical energy, 94, 155, 192 -, of energy in general, 96, 155 -, of momentum, 141 ff., 155 -, of moment of momentum, 152, 155 Conservative forces, 92, 94 -, system, 94, 98 Constrained motion, 35, 239 ff., 259 Constraints, 239 ff. -, connection with mechanical principles, 250 -, Gauss' principle of least constraint, 253 ff. Continuity, 289, 345 -, equation of, 345 ff. Continuous media, 298, 345 -, motions of, 345 Coordinates, 4 -, generalized, 391 ff. -, polar, 4, 263 -, rectangular, 4 -, spherical, 5, 249, 394 -, transformation of, 4, 249, 394 Coplanar forces, 118, 119, 204, 209 Coriolis acceleration, 222, 224, 238 Coulomb's law, 51 Couple, 204, 212 -, arm of a, 204 -, moment of a, 204 -, resultant of a force and, 209 -, theorems about, 204 Creep, 312 Crests, 320 Cross-product of vectors, 149 Current, see Electric current Curvilinear motion in a plane, 58 ff. Cycloid, 257

D'Alembert's Principle, 250, 253, 255 f., 386

Damped oscillations, 262, 265

—, kinetic energy of, 270

—, potential energy of, 270

—, total energy of, 270

Damping factor, 262, 279, 293

Damping force, 43, 262

DeBroglie waves, 402 Decay modulus, 265, 418 Deformable bodies, 298 ff. -, collisions of, 144 Degree of freedom, 386 Density, 106, 112 -, effect of variation of, on center of mass, 199 -, excess, in fluid and solid media, 362 -, mean, 361 Differential equations, additive property of solutions of linear, 318 -, linear, 37, 316, 318 —, solution of, 37, 316 Differential operator, 290, 404 Dilatational strain, 300 Dimensions, theory of, 409 Discreteness, 6 -, in problem of medium motion, 328 Displacement, 6 ff. -, angular, 11 —, complex, 273 —, velocity of, 9 —, virtual, 132 Dissipation, 260 ff. —, function, 270 Dissipative system, 260, 271 -, forced oscillations of, 273 ff. —, free oscillations of, 260, 270, 273 -, mechanical impedance of, 280 -, mechanical oscillations of, 260 ff. -, mechanical reactance of, 280 —, mechanical resistance of, 280 —, power of, 277 Distance, 5 -, units of, 5 Distributive law for vectors, 26 Divergence, of gravitational field intensity, 112 -, theorem of, 343, 347 Dot product of vectors, 25 Drag, tangential, 167 Dulong and Petit, law of, 166 Dynamical systems, 386 Dynamical theory of tides, 360 Dyne, 20, 28

Earth, motion of a particle relative to, 221 -, rectilinear motion of a particle near surface of, 34, 48 Eccentric anomaly, 81 Eccentricity of central orbits, 74, 75 Effective forces, 252 ff. Eigen-values, 405 Elastic, after effect, 312 —, bodies, 303 -, constants, 304 —, impact, 144 —, lag, 312 -, limit, 311 -, medium theory of light, 333 -, moduli, table of, 304 —, restoring force, 42, 262 —, waves, 315 ff. Elasticity, 303 —, moduli of, 305 ff. Electric circuit, 284 —, effective mass of, 284 —, impedance of, 285 -, resistance or damping factor of, Electric current, 284 Electric oscillations, 284, 285 Electromotive force, 284 Electron, 82 ff., 285 ff. -, charge on, 82, 286 -, energies in Bohr theory, 100 ff. -, energies in De Broglie theory, 402 —, free, 400 -, mass of, 83, 102, 148 -, measurement of charge on, 369 —, oscillator, 285 —, volt, 102 -, wave mechanics view of, 400 ff. Electron motion, 82 ff., 100, 152 -, Bohr's postulates for, 83 -, Bohr theory of, 82 ff., 287, 289

-, in the hydrogen atom, 82 ff.

-, quantum conditions for, 84 ff.

-, relation to planetary motion, 83

—, orbits in, 83

-, period of, 87

Electron motion, wave mechanics view of, 289, 400 ff.

Electrostatic unit of potential, 102

Element, of arc, 198

-, of area, 197

-, of surface, 200

-, of volume, 201

Elementary displacement, 9

Ellipsoid, center of mass of, 201

-, moment of inertia of, 186

-, momental, 227

-, of Poinsot, 227

Elliptic integrals, 242

Elliptical motion, 65, 74, 77, 83, 100

-, energy in, 101

Energy,

-, average kinetic, for a system of particles, 159

-, concept and potential energy, 92

-, concept, in solution of mechanical problems, 386

-, connection with work, 28

-, conservation of, 94, 96, 155, 352

-, density, in wave motion, 322

—, dissipation in oscillating systems, 270

-, electron, in Bohr theory, 100

-, -, in De Broglie theory, 402

-, equation of, 92

-, flow in wave motion, 324

-, in particle dynamics, 92 ff.

-, integral, 98, 100

-, kinetic, for a particle, 92, 159, 395

-, kinetic, for a rigid body, rotational and translational, 180, 192 ff.

-, kinetic, of system of particles, 153 ff.

-, loss of on impact, 157

-, mechanical, for particle, 93

-, molecular in a gas, 164

-, of damped oscillations, 270

—, of elastic wave in rod, 322

-, of gas, 164

-, of ground state of atom, 102

-, of particle in inverse square field,

101

-, of transverse wave in a string, 418

-, of vibrating membrane, 419

-, of wave motion in fluids, 422

-, potential, 92 ff.; see also Potential energy

-, quantum of, 288

-, radiation by atoms, 102

-, frequency distribution of, 421

Energy relations in central force field, 97

-, in oscillatory motion, 270 ff.

-, in two particle problem, 155 ff.

-, in uniform field, 96

Energy, surface, in a liquid, 373

-, total, 93

-, transfer across boundary, 422

Epoch, 40

Equation, general gas, 163

-, Kepler's, 82

-, Laplace's, 112, 412

-, of continuity in fluid motion, 345, 363

-, of equilibrium of a fluid, 340

-, of equilibrium of a particle, 119

-, of wave motion, see Wave

-, of state of a gas, 163

-, Poisson's, 112, 413

 Schrödinger's, in wave mechanics, 405

Equations, of motion, 29, 59, 141 ff.

-, canonical, of Hamilton, 407

-, component, for particle motion in a plane, 59

-, Euler's for rigid body, 228 ff.

-, Euler's hydrodynamic, 350

-, Lagrange's, 253, 393

-, of elastic medium, 316 ff.

-, of fluid medium, 349, 361

-, of motion of rigid body, 193 ff., 228 ff.

Equilibrium, 116 ff., 202 ff., 209 ff.

-, analytic conditions for, 120

—, analytic expressions of, for rigid bodies, 213

-, angle of repose for, 131

-, coplanar forces, 119, 213

-, general theorems, 120, 213

-, neutral, 345

INDEX

Earth, motion of a particle relative DeBroglie waves, 402 Decay modulus, 265, 418 to, 221 Deformable bodies, 298 ff. -, collisions of, 144 Degree of freedom, 386 Density, 106, 112 -, effect of variation of, on center of mass, 199 -, excess, in fluid and solid media, 362 -, mean, 361 Differential equations, additive property of solutions of linear, 318 -, linear, 37, 316, 318 —, solution of, 37, 316 Differential operator, 290, 404 Dilatational strain, 300 Dimensions, theory of, 409 Discreteness, 6 -, in problem of medium motion, 328 Displacement, 6 ff. -, angular, 11 —, complex, 273 -, velocity of, 9 —, virtual, 132 285 Dissipation, 260 ff. —, function, 270 Dissipative system, 260, 271 -, forced oscillations of, 273 ff. -, free oscillations of, 260, 270, 273 -, mechanical impedance of, 280 -, mechanical oscillations of, 260 ff. -, mechanical reactance of, 280 402-, mechanical resistance of, 280 —, power of, 277 Distance, 5 -, units of, 5 Distributive law for vectors, 26 -, volt, 102 Divergence, of gravitational field in--, wave mechanics view of, 400 ff. tensity, 112 Electron motion, 82 ff., 100, 152 -, theorem of, 343, 347 -, Bohr's postulates for, 83 Dot product of vectors, 25 —, Bohr theory of, 82 ff., 287, 289 Drag, tangential, 167 -, in the hydrogen atom, 82 ff. Dulong and Petit, law of, 166 -, orbits in, 83 Dynamical systems, 386 -, period of, 87 Dynamical theory of tides, 360 -, quantum conditions for, 84 ff. Dyne, 20, 28 -, relation to planetary motion, 83

-, rectilinear motion of a particle near surface of, 34, 48 Eccentric anomaly, 81 Eccentricity of central orbits, 74, 75 Effective forces, 252 ff. Eigen-values, 405 Elastic, after effect, 312 —, bodies, 303 -, constants, 304 -, impact, 144 —, lag, 312 —, limit, 311 -, medium theory of light, 333 -, moduli, table of, 304 -, restoring force, 42, 262 —, waves, 315 ff. Elasticity, 303 -, moduli of, 305 ff. Electric circuit, 284 -, effective mass of, 284 -, impedance of, 285 -, resistance or damping factor of, Electric current, 284 Electric oscillations, 284, 285 Electromotive force, 284 Electron, 82 ff., 285 ff. -, charge on, 82, 286 -, energies in Bohr theory, 100 ff. -, energies in De Broglie theory, -, free, 400 -, mass of, 83, 102, 148 -, measurement of charge on, 369 -, oscillator, 285

Electron motion, wave mechanics view of, 289, 400 ff.

Electrostatic unit of potential, 102

Element, of arc, 198

-, of area, 197

-, of surface, 200

-, of volume, 201

Elementary displacement, 9

Ellipsoid, center of mass of, 201

-e, moment of inertia of, 186

-, momental, 227

—, of Poinsot, 227

Elliptic integrals, 242

Elliptical motion, 65, 74, 77, 83, 100

-, energy in, 101

Energy,

-, average kinetic, for a system of particles, 159

-, concept and potential energy, 92

—, concept, in solution of mechanical problems, 386

-, connection with work, 28

-, conservation of, 94, 96, 155, 352

-, density, in wave motion, 322

—, dissipation in oscillating systems, 270

-, electron, in Bohr theory, 100

-, -, in De Broglie theory, 402

-, equation of, 92

-, flow in wave motion, 324

-, in particle dynamics, 92 ff.

-, integral, 98, 100

-, kinetic, for a particle, 92, 159, 395

—, kinetic, for a rigid body, rotational and translational, 180, 192 ff.

-, kinetic, of system of particles, 153 ff.

-, loss of on impact, 157

-, mechanical, for particle, 93

-, molecular in a gas, 164

-, of damped oscillations, 270

-, of elastic wave in rod, 322

—, of gas, 164

-, of ground state of atom, 102

—, of particle in inverse square field, 101

—, of transverse wave in a string, 418

-, of vibrating membrane, 419

-, of wave motion in fluids, 422

-, potential, 92 ff.; see also Potential energy

-, quantum of, 288

-, radiation by atoms, 102

-, frequency distribution of, 421

Energy relations in central force field, 97

-, in oscillatory motion, 270 ff.

-, in two particle problem, 155 ff.

-, in uniform field, 96

Energy, surface, in a liquid, 373

-, total, 93

-, transfer across boundary, 422

Epoch, 40

Equation, general gas, 163

-, Kepler's, 82

-, Laplace's, 112, 412

-, of continuity in fluid motion, 345, 363

-, of equilibrium of a fluid, 340

-, of equilibrium of a particle, 119

-, of wave motion, see Wave

-, of state of a gas, 163

—, Poisson's, 112, 413

—, Schrödinger's, in wave mechanics, 405

Equations, of motion, 29, 59, 141 ff.

-, canonical, of Hamilton, 407

-, component, for particle motion in a plane, 59

-, Euler's for rigid body, 228 ff.

-, Euler's hydrodynamic, 350

-, Lagrange's, 253, 393

-, of elastic medium, 316 ff.

-, of fluid medium, 349, 361

—, of motion of rigid body, 193 ff., 228 ff.

Equilibrium, 116 ff., 202 ff., 209 ff.

-, analytic conditions for, 120

-, analytic expressions of, for rigid bodies, 213

-, angle of repose for, 131

-, coplanar forces, 119, 213

-, general theorems, 120, 213

—, neutral, 345



Equilibrium, non-coplanar forces, 120 -, homogeneous, 346 -, incompressible, 341, 346 -, of a couple, 209 ff. -, perfect, 338 -, of a flexible string, 125 -, surface phenomena, 370 ff. -, of a fluid, 339 -, viscous, 365 ff. -, of a force and a couple, 210 ff. —, waves in a, 358 ff. -, of a ladder, 214 Fluid motion, 345 ff. —, of a particle, 116 ff. —, acceleration, 349 -, of a particle on a rough surface, -, Bernoulli's theorem, 352 ff. -, of a particle on a smooth surface, -, equation of continuity, 346, 363 -, equations of motion, 349 127 -, of a rigid body, 202 ff. —, Euler's equations, 350 -, -, conditions of, p. 213 —, steady flow in, 352 -, of a system of particles, 121 -, Torricelli's theorem, 356 -, velocity measurement in, 355 -, principle of separate, 121 Flux, gravitational, 110 —, rotational, 203, 214, 219 -, stable, 137, 344 Foot, 5 -, state of, for gas, 160 Foot pound, 28 -, translational, 203, 214 Foot poundal, 28 -, under coplanar forces, 209 Force, 17 ff. —, unstable, 137, 344 -, buoyant, of a liquid, 342 Equipotential surface, 115 Force, centripetal, 71 Erg, unit of work and energy, 28 —, damping, 43, 262 Euler's equations of motion for rigid -, definition of, 18 -, dependent on time, 52 body, 228 Euler's hydrodynamic equations, 350 —, elastic restoring force, 42, 262 Euler's theorem, use of, 406 -, frictional, 43, 130 Eulerian angles, 230 —, periodic, 273 ff. Evaporation, 372 -, static, 116 Excess density, 362 -, units of, 20 Excess pressure, 362 Force function, 99 Forced oscillations of a dissipative Fatigue, 313 system, 273 ff. Field of force, central, 66, 69 ff. Forces, central, 69 ff. -, component, 59 -, conservative, 94 -, constraint, 239, 252

—, conservative, 94
—, first power, 37 ff.
—, inverse cube, 51
—, inverse square, 44 ff.
—, inverse square attractive, 45
—, inverse square repulsive, 44
—, uniform, 33 ff.
Flexible string, equilibrium of, 125
Flow velocity gradient, 167
Fluid, 338
—, at rest, 338

-, buoyancy of a, 342

—, equilibrium of a, 339

, impulsive, 25, 52
, inverse square, 73 ff., 445
, non-conservative, 132
, non-coplanar, 120

-, coplanar, 118, 119, 204, 209

action of, 116 ff.

action of, 202 ff.

-, impressed, 251

-, equilibrium of particle under

-, equilibrium of rigid body under

Forces, polygon of, 119 —, re-action, 128 -, triangle of, 119 Fourier series, 331 Free motion, 33 ff. Free oscillations, 42 Free oscillations of a dissipative system, 260 ff. Free space, 112 Frequency, 39 -, of simple harmonic wave, 321 Friction, coefficient of, 130 —, sliding, 129 -, static, 131 -, two laws of, 129 f. Frictional forces, 129 Fundamental of vibrating string, 328 Fundamental definitions of mechan-

Galileo and the laws of motion, 17, 21 Gas, Boyle's law for, 161 —, constant, 163

—, density, 162, 164

Funicular polygon, 123

-, equation of state, 159 ff., 163

-, ideal, 160

ics, 2 ff.

b

þ,

-, kinetic energy of, 162

-, kinetic theory of ideal, 159 ff., 163

-, kinetic theory of real, 171 -, mean free path in, 166

—, molecular collisions in, 166, 170

-, molecular velocity of, 162 -, molecular weight of, 164, 166

-, perfect, see Ideal gas

—, pressure, 160

-, root mean square velocity of molecules in, 162

-, size of molecules of, 170

—, specific heat of, 163, 165

-, specific heat of, at constant pressure, 357

—, specific heat of, at constant volume, 165

-, velocity of gas stream, 357

—, virial for ideal, 160

-, viscosity of, 166 ff.

Gauss' law of normal flux, 109 ff. Gauss' principle of least constraint, 253 ff., 386 Gay-Lussac's law, 164 Generalized coordinates, 391 -, Lagrange's equations of motion in, 393 Generalized momentum, 393 Geometry of motion, 15 Gold, scattering of α-particles by nuclei of, 89 Gradient, velocity, 167 Gram, 17 Gravitation, universal, 45, 79 Gravitational attraction, 46 ff., 99 — —, in two-particle problem, 145 Gravitational constant, 20, 49 f. Gravitational field, 46 Gravitational field intensity, 106, 110 -, divergence of, 112 Gravitational flux, 110 Gravitational potential, 103 ff., 110 Gravity, acceleration of, 20, 190, 223 —, center of, 207 Gravity waves, 358 Green's theorem, 343, 347 Ground state of hydrogen atom, 102 Group velocity, 402, 408 Gyration, radius of, 182 Gyroscope, 231 -, nutation of, 233 —, precession of, 233

Hamiltonian, 407 Hamilton, canonical equations of motion, 407 —, principle of, 386 ff. -, principle of, Lagrange's equations derived from, 392 ff. Harmonic motion, see Simple harmonic motion Harmonics of stretched string, 328 -, spherical, 115, 412 Heat of a gas, on kinetic theory, 161,

Height, metacentric, 344 —, of rise in a capillary tube, 374

Helium nucleus, 87 Helmholtz resonator, 281, 295 Hereditary elasticity, 314, 315 Heredity, 314 -, coefficient of, 315 Herpolhode, 228 Homogeneous bodies, center of mass of, 196 ff. Hooke's law, 42, 269, 282, 298, 303 ff., 311, 314, 316 Horsepower, 29 -, hour, 29 Hydrodynamics, equations of, 350 Hydrogen atom, 82 -, Bohr-Rutherford model of, 82, 148 -, emission of radiation from, 102 -, energy values in Bohr model of, 101 -, excited state of, 102 -, ionization potential of, 102 -, mass of, 83, 148 -, size of, in ground state, 102 Hydrogen molecule, collisions of, 166, -, mean free path of, 170 -, size of, 170 Hydrostatic stress, 302 Hydrostatics, 338 ff. -, fundamental equations of, 340 Hyperbolic, functions, 44, 269 —, motion, 65, 74, 88 Impact, elastic, 144

—, plastic, 144
Impedance, mechanical, of oscillating system, 280
—, mechanical, resistance and reactance components, 280
—, of acoustic resonator, 282
—, of electric circuit, 285
Impressed forces, 251, 255
Impulse, 24
Inclined plane, 35, 97, 131
Incompressible fluid, 341, 346
Inductance, 284
Inertia, 2, 15

mass, 15 -, moments of, 179, 194 -, -, calculation of, 180 -, products of, 194 Inertial system, 3, 216 Initial conditions, importance of, 34, -, in inverse square field motion, Instantaneous acceleration, 12 Instantaneous angular velocity, 11 Instantaneous angular velocity of rigid body, 177 Instantaneous velocity of a particle, 9, 35 Integral equations, 315 Internal viscosity, 313 Intensity of a wave, 324 f. Invariants, adiabatic, 296 Inverse square field, 73 ff., 100 -, attractive, 73 -, energy integral for, 100 —, motion in, 73, 87, 145 ff. -, potential for, 100 -, repulsive, 87 ff. -, "time" equation, 76 Ionization potential of hydrogen, 102 Irrotational motion, 361 Isotropic medium, 316, 332 -, elastic waves in, 332

-, concept of and connection with

Joule, 28, 29 Jurin's law, 375

Kater's pendulum, 191
Kepler's equation, 82
Kepler's laws of planetary motion, 77 ff.
Kilogram, connection with pound, 17
Kilowatt, 29
Kilowatt hour, 29
Kinematics, 15 f.
Kinetic energy, 28, 154 ff.; see also
Energy
—, in collisions of spherical particles, 157

Kinetic energy, in terms of generalized coordinates, 391, of rigid body, 225 ff., 229

-, of system of particles, 153 ff.

Kinetic potential, 388 Kinetic reaction, 19

Kinetic theory, of ideal gas, 159 ff.

-, of real gas, 171

Lag, elastic, 312
Lagrange's equations of motion,
391 ff.

-, advantages of, 394

-, application of, 394 ff.

Lagrange's equations in hydrodynamics, 350

Lagrangian function, 388

-, for central force field, 395

Laplace's equation, 112, 412 Larmor's, precession, 237

-, theorem, 237

Law of areas, 71

-, of Boyle, 161

-, of Coulomb, 51, 83

-, of Dulong and Petit, 166

—, of Fermat, 421
—, of Gauss, 107 ff.

-, of Gay-Lussac, 164

--, of Hooke, 42, 269, 282, 298, 303 ff., 311, 314, 316

-, of Jurin, 375

-, of Newton, for universal gravitation, 45, 79, 223

-, of Snell, 421

-, of Stokes, for viscous fluids, 369

—, of transmissibility of pressure, 340

Laws, Kepler's, 77

-, of friction, 131

-, of motion, Newton's, 19, 254, 386 Least constraint, Gauss' principle of,

253 ff., 386

Light, corpuscular theory of, 333
—, elastic medium theory of, 333

-, electromagnetic theory of, 334

-, propagation of, 333

- law of refraction of, 421

Line of application of parallel forces, 205

Line of force, 139

Linear differential equations, 37, 316, 318

Linear strain, 299

Lines of flow in a fluid, 348

Liquids, see Fluids

Lissajous figures, 68

Loaded string, 289 ff., 419

Logarithmic decrement, 265

Lorentz-Einstein transformation

equations, 401

Lost forces, 251

Magnet, 51, 139, 414

Mass, 2, 15 ff.

-, additive property of, 16

-, center of, 141 f., 147, 152, 207

-, definition of, 16

-, measurement of, by balance, 208

-, of electron, 83, 148

-, of planets, 79

-, reduced, 146

—, units of, 16

-, variation of, with velocity, 403

Material particle, definition of, 2

Mean anomaly, 82

Mean free path in gas, 166

Mean solar second, 5

Mechanical energy, 93, 96, 155

—, conservation of, 94, 96, 155, 192;

see Energy

Mechanical equivalent of heat, 165

Mechanics, definition of, 1

-, elementary concepts of, 1

-, fundamental principles of, 2, 19, 250

-, of a rigid body, 175 ff.

—, of fluids, 338 ff.

-, wave, 289, 399 ff.

Membrane, motion of, 418 f.

Metacenter, 344

Metacentric height, 344

Meter, 5

Method of components for equilibrium problems, 214, 216

Minimum principle in mechanics, 387 Moduli, elastic, 305 ff.

elastic, experimental determination of, 308

-, elastic, relations among, 305

-, elastic, table of, 304

Modulus, of complex number, 263

-, of volume elasticity, 305

-, shear, 305

—, Young's, 305, 307, 316, 323 f.

Molecular weights, 164

Molecules, 159 f., 286

Moment, 150

—, of a couple, 204

—, of a magnet, 414

—, of acceleration vector, 150

—, of momentum, 149, 151, 193; see Angular momentum.

Moment of inertia, 179, 194

 —, calculation of, for homogeneous elliptical disc, 183

-, calculation of, for homogeneous solid rod, 181

-, calculation of, for sphere, 184

-, components of, 194

-, table of, 186

Momental ellipsoid, 227

Momentary forces, 25

Momentum, 19

—, angular, 179, 193, 216, 218, 225, 228, 230

-, connection with force and impulse, 24

-, conservation of, 143

-, moment of, 149, 151, 193

—, resultant, 143

Motion, angular, 11

—, along an inclined plane, 35 —, central field, 66, 69 ff.

-, circular, 11

-, components of, 58

-, components of, in a plane, 58

—, constrained, 35, 239 ff., 259

-, curvilinear in a plane, 58 ff.

-, definition of, 2

equations of, see Equations of motion

—, free, 33

-, in a field directly proportional to the distance, 37 ff.

—, in a repulsive inverse square field, 44, 87 ff.

—, in a straight line, 33

-, in a uniform field, 33

—, in an attractive field, 45

—, in an inverse square field, 44, 73, 145

—, independence of translational and rotational, 191, 195

-, laws of (Newton), 19, 254, 386

-, linear, 11

-, of a conservative system, 261, 270

-, of a fluid, 345 ff.

-, of a freely falling particle, 34

—, of a particle on earth's surface, 221

—, of a particle on a smooth inclined plane, 252

-, of a pendulum, 187, 239, 269

-, of a rigid body, 153, 175 ff., 216

—, of a spring, 41 ff.

—, of a string, 325

-, of a system of particles, 141 ff., 386 ff.

—, of a top, 229 ff.

—, of electrons, 82 ff., 152

-, of elastic medium, 315 ff.

-, plane curvilinear, 58 ff.

—, planetary, 77 ff., 100, 145, 152

—, rectilinear in resisting medium, 53 ff.

—, relative to moving axes, 216 ff., 219 ff.

-, rotational, 175

-, steady, of a fluid, 352 ff.

-, translational, 175

-, turbulent, of a fluid, 366

-, with constant acceleration, 33

Neutral equilibrium, 345

Neutron, 286

Newton meter, 29

Newton's law of universal gravitation, 45, 79, 223

Newton's laws of motion, 19, 254, 386 Nodes, 329 Non-conservative forces, 132 Non-conservative systems, 96; Dissipative system Non-coplanar forces, equilibrium of, 120 -, parallel, 205 Non-homogeneous bodies, 199 -, center of mass of, 199 ff. Non-uniform strain, 299 Normal thrust, 128 Nuclear physics, 153 Nucleus, atomic, 82 —, charge on, 82, 87 -, helium, 87 -, mass of, 148 Nutation, 233 One-to-one correspondence, 5 Operational calculus, 290, 294 Operators, 290, 404

Oscillations, 259 ff. —, acoustic, 281 -, atomic, 286 f. —, damped, 262, 265 ff. -, electrical, 284 f. —, energy of, 270 —, forced, 273, 275 -, free, 42, 260 ff., 270 ff., 273 —, natural, 273 -, of a resonator, 281 -, of a spring, 42, 260 —, of a string, 325 -, of a system with one degree of freedom, 260 ff. -, of a system with several degrees of freedom, 289 ff. —, of charged particles, 286 -, periodic, 293 —, undamped, 37 ff., 259, 271 Oscillograph, 68 Oscillator in atomic theory, 285

Parallel axes, theorem of, 182
Parallel forces, 204 ff.

— —, line of application of, 205, 206

Particle, 2 —, acceleration of, 12 -, energy of, 28, 92 ff. Particle equilibrium, 116 ff. ---, on rough surface, 129 ---, on smooth surface, 127 — —, principle of virtual work in, 132 Particle, motion of, on earth's surface, 221 Particle, motion of, see Motion -, position in time and space, 3 -, rectilinear motion of, 33 ff. -, velocity of, 9 Particles, system of, 141 ff. —, energy of system of, 153 ff. -, motion of system of, 141 ff. Pascal's principle, 339, 340 Pendulum motion, 187, 239, 269 -, physical significance of, 190 Perception, modes of, 3 Perfect fluid, 338; see also Fluid Perfect gas equation, 163 Perihelion in elliptical motion, 78 -, motion of, in relativity, 410 Period of planetary motion, 77, 79 -, of simple harmonic motion, 39 -, of simple harmonic wave, 321 Periodic motion, see Oscillations Permanent set, 312 Phase, 40, 322 —, difference, 274, 365 Physical pendulum, 187 ff. -, center of oscillation of, 189 —, equation of motion of, 187 -, frequency of, 188 —, use of, in determination of g, 190 Physical theory, nature of, 1 Pitot tube, 355 Planck's constant, 84, 101 f., 288, 402 Plane, curvilinear motion in, 58 ff. -, equilibrium of forces in a, 118, 119, 215 -, harmonic waves, 319, 363 ff. —, motion in a, 58 ff.

-, waves in a fluid, 363

—, waves in a rod, 319

Parametric equations of motion, 61

INDEX

Planetary motion, 77 ff., 100 f., 145 f., 152	—, equation, 245
	—, units of, 29
—, Kepler's laws of, 77 —, period of, 77, 79	Power factor of an electrical circuit 295
-, "time" equation in, 80	Precession, Larmor, 237
Planets, determination of mass of,	—, of a gyroscope, 230 f.
79	
Plastic bodies, 303	—, of a top, 233 Precessional velocity, 230, 231
—, impact, 144	Pressure, 160, 302, 339 f.
Plasticity, 304	—, dependence on depth in a fluid,
Poinsot's ellipsoid, 227	341
Poiseuille's formula, 368	
Poisson's equation, 112, 115, 413	—, excess in a fluid, 362
Poisson's ratio, 300, 307	—, inside liquid in shape of sphere, 381
Polhode, 228	—, inside soap bubble, 381
Polygon of forces, 119	- of a gas on kinetic theory 100 g
Position, concept of, 3	-, of a gas, on kinetic theory, 160 ff.
—, coordinate representation of, 4	—, Pascal's principle concerning, 339—, transmissibility of, 340
—, relativity of, 3	Primary inertial system, 3, 216
—, vector, 4, 8	Principal axis, 227
Potential, as solution of Laplace's	Principle, Archimedes', 341 f.
equation, 113, 115	—, D'Alembert's, 250 ff., 386
—, calculation of, 103 ff.	—, energy, 94
-, definition of, 99	—, Fermat's, 421
-, due to charge distribution, 412 f.	C , , , ,
-, due to spherical shell, 103 ff.	—, Gauss' (of least constraint), 253 ff.
—, electrostatic, 99	—, Hamilton's, 386 ff.
—, gravitational, 99	-, of conservation of momentum,
-, ionization, of hydrogen, 102	141 ff., 143, 417
-, magnetostatic, 99	-, of independence of translational
Potential energy, 92 ff.	and rotational motions of a rigid
-, associated with elastic wave,	body, 191, 195
323 f.	—, Pascal's, 339 f.
—, in central force field, 98	-, of superposition, 293
—, in uniform field, 96	—, of virtual work, 132
—, of a particle, 93, 98	Principles, of hydrostatics, 338 ff.
—, of a system of particles, 137, 154 f.	-, of mechanics, connection with
—, of fluid, due to pressure, 352	constraints, 250
-, of fluid, due to external forces,	-, of relativity, 400
351	Products of inertia, 194
—, of simple harmonic oscillator, 270	Projectile motion, 60
—, of vibrating string, 418	-, component velocities in, 61
Pound, of force, 20, 29	—, in a resisting medium, 62
—, unit of mass, 17	-, parametric equations of, 61
—, weight, 20	—, range, 61
Poundal, 20	Proton, 286
Power, 29	Public space and time, 3

Quadric surface, 227

Quantization, of electron motions in hydrogen atom, 84

- -, of electron motions in oscillator, 287, 289
- -, Schrödinger's method of, 404 Quantum, conditions, 84, 288, 400
- -, numbers, 84, 101
- ___, of energy, 288
- -, theory of atomic structure, 82 ff.
- -, theory of radiation, 102
- Quantum theory equation, 402

Radian, 11

Radiation, absorption of, by oscillators, 287

- -, acoustic, 283
- distribution of energy in electromagnetic, 421
- —, from an atom, quantum theory, 102
- -, resistance, 283

Radium C, 89

Radius, of gyration, 182

—, vector, 69

Range, 61

Rarefaction, 364

Reaction, kinetic, 19

-, of a surface, 128

Recoil (of a gun), 21

Rectilinear motion in a resisting medium, 53

Reduced mass for system of particles,

Reference system, 3, 30

Reflection of sound, 422

Relative motion, 216, 219

Relativity, theory of, in wave mechanics, 400

-, motion of perihelion of Mercury, 410

Relaxation time, 269, 312

Repose, angle of, 131

Resistance, acoustic, see Impedance

-, electrical of a circuit, 285

- -, of a plane, 131
- -, of medium to motion of a body, 57

—, passive, 132

Resonance, 276 ff., 287

- -, effect of damping on, 276 ff., 287
- -, frequency, 276
- -, maximum dissipation at, 277
- -, of electrical circuit, 285
- -, of mechanical oscillator, 276
- -, sharpness of, 279

Restitution, coefficient of, 144, 145,

Restoring force, 262, 290

Resultant, magnitude of, for vectors in general, 7

- —, moment, 203, 213
- -, momentum, 143
- -, of forces in a plane, 209 ff.
- -, of parallel forces, 205
- -, of parallel non-coplanar forces, 205

Reversible pendulum, 191

Rigid body, acted on by parallel forces, 204 ff.

Rigid body, angular acceleration of,

- -, angular momentum of, 179, 193
- -, angular velocity of, 177
- -, axes fixed in, 216 ff., 219 ff., 228
- —, center of mass of, 196 ff.
- -, center of gravity, 202 ff.
- —, conditions for equilibrium of, 203, 213
- —, definition of, 175
- -, displacement of, 175 f.
- —, equilibrium of, 202, 209
- —, general equations of motion of, 193
- -, kinetic energy of, 180, 226
- -, mechanics of, 115 ff.
- -, moment of inertia of, 179 ff., 194
- -, motion of, 175 ff., 216 ff.
- -, motion of center of mass of, 195 ff.
- -, moving axes and, 216, 219, 228
- -, plane motion of, 191 ff.
- -, rotation of, 175 ff., 225
- -, rotation of, about fixed axis, 178 ff.
- -, rotational equilibrium of, 203, 219

Rigid body, translation of, 175 f.

-, translational equilibrium of, 203

—, work-kinetic energy relation for rigid body with one point fixed, 226 Rigidity, 305; see Shear modulus Rotation, and angular velocity, 176 ff.

-, equilibrium of rigid body with

respect to, 203, 219

-, kinetic energy of, 180, 192, 225 ff.

-, of axes, 216

, of rigid body about a fixed axis,178 ff.

—, of rigid body about a fixed point, 175, 225

—, of a top or gyroscope, 229 ff.

-, work-kinetic energy theorem for, 180

-, simple harmonic motion of, 187 f., 310

Rough surface, equilibrium of a particle on, 129 ff.

-, resistance of, 131

Satellite, mass of planet determined by motion of, 79 Scalar product of two vectors, 26

Schrödinger treatment of electron motion, 403 ff.

—, wave equation, 405

Second, mean solar, 5
Series, expansion for period

Series, expansion for period of pendulum, 242

—, Fourier, 331

Shear, measurement of, 300

-, modulus of, 305

-, finding of modulus of, 308 ff.

Shearing, strain, 300 ff.

-, in liquids, 301

—, stress, 303

Ship stability, 344

Simple harmonic motion, 39 ff., 240

-, amplitude, 39, 265

-, composition of, in a plane, 64 ff.

-, damping force, 43, 262

—, energy equation for, 95

-, epoch, 40

-, frequency, 39, 265

—, kinetic energy in, 270

-, of an electron, 287 f.

—, of a spring, 42, 260

-, parametric equations of motion, 64

—, period, 39

—, phase, 40, 265

—, potential energy in, 270

-, quantization of, 287, 403

Simple pendulum, 189, 239, 260, 269 Sleeping top, 233

Smooth surface, equilibrium on, 127

, motion of particle on, 243, normal thrust on, 128

Soap bubble, pressure inside, 381 Solid angle, 109

Solutions, adsorption, 382

—, surface tension, 382

Sound waves, see Acoustic waves Space, 3

Specific heat of gas, 163, 165

-, connection with sound velocity, 365

-, constant pressure, 357

-, constant volume, 165

-, definition, 165

—, law of Dulong and Petit, 166 Spectrum of radiation from hydrogen gas, 102

Speed, 9

Sphere, moment of inertia of, 184

-, potential of, 106

—, pressure inside liquid in shape of, 381

Spherical, coordinates, 394

-, harmonics, 115

Spherical shell, potential due to, 103 ff.

Spherical particles, 144, 156 Spherical waves, 363, 422

Stable equilibrium, 137, 344

State of a dynamical system, 386 Static forces, 116 ff.

_, equilibrium of, 116 ff., 202 ff.

Static friction, 129

Statics of a particle, 116 ff.

-, of a rigid body, 202 ff.

Steady flow of a fluid, 352 Steady state solution of oscillation problems, 273

Stiffness, coefficient, 261

—, of a spring, 42 Stokes' law, 369

Strain, 298 ff.

-, dilatational, 300

_, lateral, 300

_, linear, 299, 305

-, non-uniform, 299

-, shearing, 300

—, uniform, 299

-, volume, 298, 304

Stream lines, 348, 353, 366

Stress, compressive, 301

-, hydrostatic, 302

-, shearing, 303

-, tensile, 303, 305

String, transverse vibrations in, 325 ff.

Subtraction of vectors, 7

Superposition principle, 10, 16, 293

Surface, energy, 373
—, equipotential, 115

-, integral, 109, 342, 343, 347

-, phenomena, 370

—, rough, 129

-, smooth, 127, 243

Surface tension, definition of, 371

—, measurement of, 371, 373, 375, 381 f.

Surface waves in a liquid, 358

Suspension bridge, 124

System, conservative, 94, 98, 271

-, coordinate, 4

—, dissipative, 260, 271, 273

—, mechanical, 386 —, reference, 3, 30

System of particles, 141 ff.

-, center of mass of, 142, 152 ff.

-, conservation of energy in, 155

—, conservation of momentum in, 143 ff.

-, conservation of moment of momentum in, 152

—, energy of, 153 ff.

—, equations of motion of, 141 ff., 149. 153

-, equilibrium of, 121

-, kinetic energy of, 153 ff.

-, potential energy of, 137, 154 ff.

—, motion of, 141 ff., 153, 387

-, moment of momentum of, 151

-, resultant momentum of, 143

-, virtual work as applied to, 185

-, work-energy theorem for, 154

Tangential viscous drag, 167

Tensile stress, 303

Tension in a string, 117, 125, 126, 325

Theorem of parallel axes, 182

Thrust, normal, 128

Tides, dynamical theory of, 360

Time, 3

-, absolute, 6

-, average of kinetic and potential energy, 271

-, continuity of, 6

-, public, 5

-, relativity of, 6

-, unit of, 5

Time equation, 72, 76, 244

Top, 229 ff.

-, energy equation of, 231

-, equation of motion of, 231

-, motion of, 229, 233

-, nutation of, 233

-, precessional velocity, 231 f.

—, spin velocity, 231, 233

Torque, 149 ff.

Torricelli's theorem, 356

Torsion, balance, 49

-, of a cylinder, 309

—, of a wire, 314

Transfer of momentum, 168

Transient oscillation, 273

Translation, motion of, for rigid body, 175, 191

Translational equilibrium of rigid body, 203

Transmissibility of pressure, law of,

Transverse waves, 325, 332

Triangle of forces, 119

Trough in wave motion, 320

Tubes of flow, 348

Tuning fork, as dissipative system. 265

Turbulent motion of a fluid, 366 Two particle problem, 145 ff.

—, atomic structure, 148

-, center of mass, 147

-, energy relations, 155 ff.

-, equations of motion, 145

—, planetary motion, 145

—, reduced mass, 146

-, with gravitational attraction, 145

-, virtual work and, 135

Undamped oscillator, 271, 288 Uniform field, 33 ff.

-, energy relations in, 96

-, motion in, 33

Uniform motion, 19

Unit vector, 26, 150

Units, dimensions of, 409

—, English, 5, 17, 20 —, metric, 5, 17, 20

-, of angular displacement, 11

-, of energy, 28

—, of force, 20

-, of length, 5 —, of mass, 16

-, of power, 29

—, of time, 5

-, of velocity, 9

-, of work, 28

Unstable equilibrium, 137, 344

Van der Waals' equation of state, 172 Variation of integral in Hamilton's principle, 387

Vector, 4

—, acceleration, 59

-, analysis, 10

-, difference, 7

-, equation, 7

—, polygon, 7

-, resolution, 8

Vectors, addition of, 7

-, commutative law for, 25

-, cross product of, 149, 219

-, definition of, 4

-, distributive law for vector cross product, 149

-, distributive law for vector dot product, 26

—, physical significance of, 4

-, position, 4

-, dot product of, 25

-, rectangular components of, 5

-, resultant, 7

-, subtraction of, 7

-, unit, 150

Velocity, angular, 11, 221

-, average angular, 11

-, components of, 9

-, critical, in viscous fluid, 366

—, displacement, 6

—, gradient, 167 -, group, 402, 408

-, instantaneous, 9

—, of waves, 318

—, potential, 361

-, precessional, 230 f.

-, relative to moving axes, 216 ff.

—, root mean square, 162

-, uniform angular, 40 -, uniform for a particle, 33

-, units of, 9

-, vector representation of uniform

angular, 12

Velocity potential, 361

Vena Contracta, 356

Venturi water meter, 354

Vibrations, see Oscillations Virial, of Clausius, 157 ff.

-, for ideal gas, 160

-, for interaction forces, 171 ff.

—, theorem, 159

Virtual displacement, 132

—, work, 132 ff., 135, 138

Viscosity, coefficient of, 167, 366

-, determination of, for liquids, 368 Viscosity of gas, 166

Viscous fluids, 301, 304, 313, 365 ff.

-, critical velocity in, 366

Volume, elasticity, 282, 305

-, strain, 298

-, stress, 302

Water meter, Venturi, 354

Watt, 29 Wave, equation, 360, 363, 405

-, group, 402

-, mechanics, 289, 399 ff.

a, motion, 315 ff., 319, 363

Waves, acoustic, 360, 363, 368

-, compressional, 358, 360

-, DeBroglie, 402

-, elastic, 319

-, elastic in a solid rod, 322

-, gravity, 358

-, in fluids, 358 ff.

-, intensity of, 324

-, longitudinal, 319, 331

-, plane, 319, 363

-, plane harmonic, 363 f.

-, simple harmonic, 320 ff.

 , space average of kinetic energy density, 323 space average of potential energy density, 323

-, spherical, 363

-, stationary in stretched string, 327

-, surface, 358

-, surface tension, 358

—, time average of kinetic energy density, 322

-, transverse in a string, 325, 332

—, types of elastic waves in solids, 331 ff.

-, velocity of, 318

Weight, 20, 205 f., 208

Work, 27, 92

-, connection with energy, 28, 92

-, unit of, 28

-, virtual, 132 ff., 135, 138

Work-energy theorem, 28, 92, 154, 180

Yield point, 311

Young's modulus, 305, 307, 316, 323

-, evaluation of, 308



LONGMANS, GREEN AND CO. LTD.

6 & 7 CLIFFORD STREET, LONDON, W.I

NICOL ROAD, BOMBAY, I

17 CHITTARANJAN AVENUE, CALCUTTA, 13

36A MOUNT ROAD, MADRAS, 2

LONGMANS, GREEN AND CO. INC.

55 FIFTH AVENUE, NEW YORK, 3

LONGMANS, GREEN AND CO.

215 VICTORIA STREET, TORONTO, I

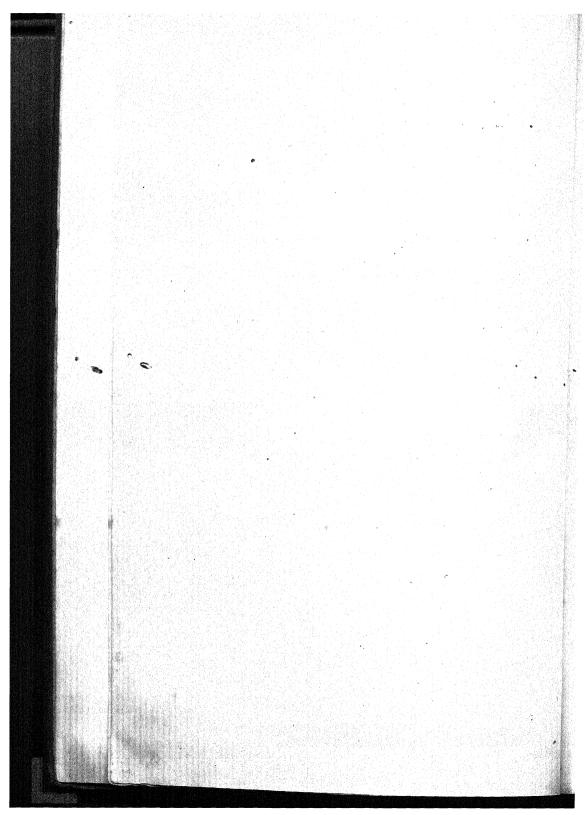
First Published - 1942
Second Edition - 1945
Reprinted with Appendix 1946 and 1948

CODE NUMBER: 85387

PRINTED IN GREAT BRITAIN BY
LOWF AND BRYDONE PRINTERS LIMITED, LONDON, N.W. 10

DEDICATED TO MY WIFE AND CHILD





FOREWORD

Dr. Tolansky's book has its origin in a series of lectures which for many years has formed part of the course given to the second year Honours Physics students at Manchester University. An undergraduate must spend most of his time in mastering the classical treatment of the main divisions of physics-heat, including thermodynamics and kinetic theory, optics, electricity and magnetism. In my opinion the teacher must restrain himself from introducing too much "modern physics" into an Honours course. When all is so new it is hard to preserve a sense of proportion. Much of the story of modern physics is that of the triumphant discovery of new facts by means of apparatus on a previously undreamed-of scale which is often a triumph of engineering design and construction. The fitting of these facts into a clear pattern has either not yet been achieved, or is only to be apprehended by the advanced expert; and therefore their study does not form a good training ground for the undergraduate.

At some point, however, the student has to be introduced to the fundamental change in outlook which has taken place since the beginning of this century. He will wish to understand where classical physics breaks down, and why heroic measures had to be adopted in discarding old ideas. He should begin to appreciate the power of the new quantum mechanics, and something of the symmetry and beauty of its conceptions. In order to achieve this result it is not necessary to give him a recital of the latest discoveries. A treatment which is largely historical forms a good basis, because it is so important that the order should be right, that he should recapture something of the bewilderment of the classical physicist, and then his excitement when new conceptions of the physical world resolved chaos into pattern.

The course at Manchester was the result of a belief that the student should get this introduction in his second rather than his third year. It formed a background for his final year of intensive study. At the same time, as he could not be overloaded with detail at this stage in his studies, difficult mathematical treatments were avoided and the main stress was on qualitative description of phenomena and their interpretation.

This course became a tradition in the Manchester Physics Laboratory, and many members of the staff contributed to it. It has been attended by quite a number of Ph.D. candidates from the chemical and other departments. It has always been one of the most interesting courses for the teacher to plan, and

I think the students have found it interesting too.

A book cannot take the place of lectures. Particularly in a course such as this, when the lecturer is trying to give a new trend to his students' thoughts, he must depend on numerous graphic illustrations, on colloquialisms, on attacks from a new angle when he sees that his class has not grasped the point, and other devices which are the justification of the lecture system but out of place or impossible in a book. Yet the student must have a book to which he can refer, and there has been a real need for a first introduction to atomic physics which will give him what he wants without taking him too far. In my opinion, Dr. Tolansky has admirably supplied this need. He gives the right emphasis to each part of the subject, and the standard is what it should be. I believe his book will be warmly welcomed by students and teachers, and I wish it every success.

W. L. BRAGG.

January, 1942.

PREFACE TO THE SECOND EDITION

In preparing this new edition the opportunity has been taken to check the whole of the text carefully so as to eliminate errors which occurred in the first edition and to improve and clarify the phraseology in many places. I am particularly grateful to Dr. N. Feather of Cambridge for his help in reading the whole book with critical care.

Apart from these corrections the book as a whole remains practically unaltered, for there has been little progress in fundamental atomic physics since the writing of the earlier edition. A few minor alterations have however been made. The more important amongst these include, the recent modification of Townsend's theory of ionisation by positive ions, a more satisfactory derivation of Langmuir's thermionic equation, the removal of some obscurity in spectroscopic notation, alterations to six diagrams, and the bringing up to date of the numerical values of the fundamental constants, which are of course always being slightly modified.

My thanks are again due to Dr. W. H. Taylor of the College of Technology, Manchester, for a further revision of the chapter dealing with crystal structure.

S. TOLANSKY.

This book has been written in an attempt to meet a real need. It aims at introducing in an elementary manner a fairly comprehensive survey of modern atomic physics to students who have already passed through an Intermediate course. It is based upon a lecture course which the writer has delivered for some years past to Second Year Honours (Physics) and Third Year Honours (General) students at the University of Manchester.

Every effort has been made to keep up to date the rapidly expanding material. With so wide a field to cover selection of subjects is inevitable, and as a result omissions arise of necessity. It is felt that the inclusion of chapters upon magnetism, electron optics, molecular spectra, low temperature, and many other aspects all of which can legitimately be classed under the heading of atomic physics, would add so much to the length of the book as to render it too cumbersome for a course extending over one year only.

An extensive list of references has not been included since it is unlikely that the student for whom this book is intended will find time to go to the very large number of original sources involved in so wide a field

The writer wishes to express his deepest appreciation to Sir Lawrence Bragg, F.R.S., Cavendish Professor of Physics in the University of Cambridge, who has been of the greatest help and has given invaluable advice. He has been kind enough to revise many chapters. The writer has had the further privilege of being able to refer freely to the lecture notes of Sir Lawrence Bragg.

The writer's warmest thanks are due to Professor P. M. S. Blackett, F.R.S., Langworthy Professor of Physics in the University of Manchester, who has given him constant encouragement and help in both the selection of material and revision. Dr. Janossy of Manchester University and Dr. W. H. Taylor of the College of Technology, Manchester, have kindly assisted by revising chapters. Dr. N. Feather of Cambridge has been particularly helpful in revision, and his assistance has been of great value.

January, 1942.

S. TOLANSKY.

CONTENTS

CHAP.	in the contract of the contrac	PAGE
1	THE CONDUCTION OF ELECTRICITY THROUGH	
	GASES	1
2	THE ELECTRIC DISCHARGE THROUGH GASES AT	
	Low Pressures	26
3	THE CHARGE AND MASS OF THE ELECTRON .	41
4	Positive Rays and the Mass Spectrograph	54
5	THE QUANTUM THEORY	75
6	THE THERMIONIC EFFECT	95
7	PHOTO-ELECTRICITY	114
8	ATOMIC SPECTRA AND THE PERIODIC TABLE .	131
-9	X-RAYS AND THEIR PROPERTIES	151
10	THE STRUCTURES OF CRYSTALS	167
11	THE WAVE MECHANICS	188
12	ELECTRON COLLISIONS IN GASES	203
13	RADIOACTIVITY AND THE RADIOACTIVE TRANS-	
	FORMATIONS	210
14	THE a-PARTICLE	227
15	The β -rays	249
16	The γ -rays	265
17	COSMIC RADIATION: THE POSITRON AND THE	
	Meson	273
18	THE NUCLEUS OF THE ATOM	295
19	THE RELATIVITY THEORY	318
	APPENDIX	336
	INDEX	341



CHAPTER 1

THE CONDUCTION OF ELECTRICITY THROUGH GASES

Introduction

The rapid development of modern atomic physics since the close of the nineteenth century has been due largely to the study of electrical discharges, radioactivity, and the properties of electromagnetic radiation in its various forms. As far back as 1803, Dalton showed that matter is atomic in nature, and ten years later Prout suggested that the different kinds of atoms are built up out of the simplest, hydrogen. The discovery of the Periodic Table of the elements by Mendeleeff pointed to the existence of some recurrent structure factor within the atoms themselves, hinting at the probable existence of sub-atomic constituent components. In the hands of Sir J. J. Thomson, the study of the discharge of electricity through rarefied gases led to the discovery of the electron, and it became clear that electricity is atomic in nature also, an electric current or an electric discharge consisting of the flow of negatively charged atoms of electricity, each having a mass about two thousand times less than that of the hydrogen atom. These particles, the electrons, were found to be sub-constituents of every kind of atom. They carry a charge which is the smallest that can be observed in nature, all charges being simple multiples of the elementary unit, the electronic charge.

Soon after the discovery of the electron, Rutherford's researches in the new science of radioactivity led him to conclude that an atom is formed out of a massive small nucleus around which rotate the relatively light electrons. More recent work has revealed the fact that the individual atomic nuclei are highly complex, being built up by different rearrangements of a few fundamental particles, not so very different in fact from the structure suggested in a crude fashion so long ago by

Prout. Some of these atomic nuclei are unstable and occasionally spontaneously disintegrate with the emission of radiation. This is the essential phenomenon of radioactivity.

The knowledge that both matter and electricity are atomic prepared the way for the development of Planck's quantum theory, which proved that radiation is also atomic. According to the quantum theory, as elaborated by Einstein, electromagnetic radiation travels in small discrete packets, called quanta, which are in effect atoms of radiation. Electromagnetic radiation is no longer to be thought of as a continuously expanding attenuating wave form, but as a stream of particles, the energy of each remaining constant whilst it travels outwards with the velocity of light. Thus the classical concepts of continuity which flourished at the time of Maxwell and Kelvin have now been discarded in favour of atomicity conceptions. Most of the important modern developments began with the study of the discharge of electricity through gases. This will now be discussed.

The electrical conductivity of gases

A gas in its normal state hardly conducts electricity at all, a fact known for centuries. However, when a carefully insulated gold leaf electroscope is charged, even if precautions are taken to eliminate leakage from the gold leaf supports, a slow discharge of electricity is observed. Experiments conducted by C. T. R. Wilson on this effect in 1900 proved that the leakage was taking place through the air contained in the electroscope chamber. The observed conductivity was very small and was proportional to the amount of air enclosed in the apparatus. Normal air is, in fact, almost a perfect insulator, and indeed were this not so the study of electrical phenomena might never have developed at all.

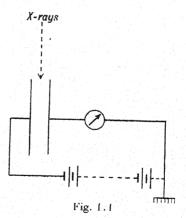
Although the conductivity of gases in the normal state is very small indeed all gases can be put into a relatively highly conducting state by a number of means. Thus, for instance, the combustion gases drawn from above flames are founds to be conducting. Conductivity can be produced if a gas is irradiated by ultra-violet light or by X-rays, or if electric sparks are passed in it, and so on. The state of conductivity, once created, exists for some time, but if not continuously

renewed keeps on diminishing until it disappears. This can be proved by first producing a conducting state in a gas and then sucking this gas through a long tube into an electroscope. The electroscope is discharged, showing that the gas retains its conductivity during its passage through the long tube agent producing the state of conductivity is removed the power of the gas to discharge the electroscope quickly diminishes and vanishes. When the conducting gas is passed through a cotton wool plug or bubbled through water the conductivity disappears. The same occurs if it is passed between a pair of plates maintained at a large potential difference. properties prove that the conductivity is due to the presence of electrically charged particles in the gas. Observations show that a gas in a conducting state is not charged as a whole, hence an equal number of negatively and positively charged particles must exist in this state. The particles are called negative and positive ions respectively, the production of conductivity being called ionisation.

Variation of current with applied voltage

·If a voltage is applied across two plates between which there is an ionised gas, an electric current flows, since the

charged particles move towards the plates. When the potential difference between the plates is varied, the resulting ionisation current does not obey Ohm's law. The variation of current with potential difference can be studied with the simple apparatus shown in Fig. 1.1. X-rays act as a source producing a constant number of ions per second between the plates. An electrical field is maintained across these by



means of a large battery. When the current passing between the plates is plotted against the applied voltage, the curve shown in Fig. 1.2 is obtained. The current at first obeys Ohm's law but then quickly ceases to increase, reaching a

constant value called the saturation current. When a very high potential is applied the current suddenly begins to increase very rapidly. For air at S.T.P. the ionisation saturation current remains constant up to an applied field of 30,000 volts

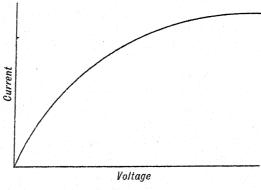


Fig. 1.2

per centimetre before suddenly rising. The value of the saturation current depends only upon the amount of ionisation between the plates. It therefore increases with increasing volume and increasing pressure between the plates. It is proportional to the mass of gas between them.

Recombination of ions

It is clear that when saturation is reached the number of ions created by the ionising agent is equal to the number removed by the applied field. However, even when no field is applied to remove the ions created, the number present, for a given agent, does not go on increasing indefinitely but after some time reaches a steady state. This is occasioned by the fact that the positive and negative ions tend to recombine when they are brought into collision with each other as a result of gas kinetic movements. The recombination of a pair of oppositely charged ions leads to the disappearance of two charges so that recombination destroys the existence of the ions. Consider what takes place if we have a constant ionising agent producing q positive and q negative ions per second in the gas, each with an electrical charge e. At a

given time let there be n ions of each kind in a cubic centimetre of the gas. The probability of a positive ion colliding with a given negative ion is proportional to n, and as there are n positive ions in the cubic centimetre of gas, the number of collisions between positive and negative ions is proportional to n^2 , i.e. it will equal an^2 where a is a constant called the coefficient of recombination. The rate of increase of ions after time t is given by $dn/dt = q - an^2$. If the ionising agent is removed, q is made zero and we have $dn/dt = -an^2$. When integrated this gives $n = n_0/(1 + n_0 at)$, n_0 being the value of n at time t = 0.

The recombination coefficient a was first measured by Rutherford by the application of the formula just derived for n. The apparatus used was simple. Ionised air was sucked through a tube at a known rate. The ionisation at three points along the axis was measured by means of the electrodes introduced for the purpose. The saturation ionisation current at any point is proportional to the number of ions there. Since the velocity of the air current can be measured, the time taken to pass between points a known distance apart can be determined. Thus the formula $n = n_0/(1 + n_0 at)$ can be applied and a value for a derived. The tube is wide in order to reduce ion losses by diffusion to the walls.

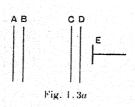
Since Rutherford's first measurements, many ingenious and more accurate methods for measuring a have been devised, and values for the recombination coefficient are known for many gases. Most of these methods require a knowledge of the velocities acquired by the ions in electric fields. The study of ionic velocities is of practical importance in connection with many aspects of the electrical discharge through gases and furnishes valuable information about the ultimate nature of the ions. The next section will therefore deal with the measurement of ionic velocities.

The mobilities of ions

When gaseous ions are produced between two plates maintained at a potential difference they accelerate towards the plates under the influence of the field. If the field strength is X and if e and m are respectively the charge and mass of each ion, the acceleration is a=Xe/m. After a short time, t, each ion collides with a neutral gas molecule and is effectively

brought to rest. The time t is that needed to travel over the mean free path between the gas molecules and can be calculated by the aid of the kinetic theory of gases. The ion, starting from rest, is again accelerated by the field and once more is brought to rest after time t. The velocity acquired at the end of each of these mean free paths is v'=at=Xet/m. Clearly, as the collisions are repeated right through the gas, the ions travel across with a net velocity and not with an acceleration. The mean value of this drift velocity is v=Xet/2m, since the ions begin from rest after each collision. For any particular type of ion in a gas at a given pressure, e, m, and t are constants, hence the ionic velocity is v=kX where k is a constant and is equal to the velocity acquired by the ion due to the action of a field of 1 volt per centimetre (i.e. X=1). This constant, k, is called the ionic mobility.

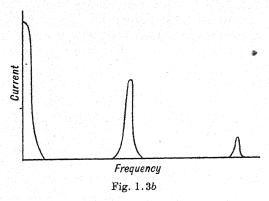
There are many different ways for measuring ionic mobilities, not all of which are equally good. The earliest measurements were made in 1897 by Rutherford, and since that time the mobilities of both the positive and negative ions in a large number of gases and vapours have been determined. We shall consider here only one of the most recent methods, due



to Tyndall. This is capable of a high order of accuracy and is illustrated in Fig. 1.3a. A, B, C, D are four similar sheets of metal gauze and E a plate connected to an electrometer. Ions are brought up to A by means of a suitably directed field. An alternating potential is applied between A and B so that

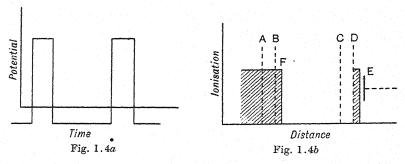
successive bursts of ions are pulled through into the space BC. The main field is maintained between the gauzes B and C, the ions travelling to C in a time depending upon the mobility. An alternating potential in phase with that between A and B is also applied between C and D. For certain values of the frequency of the alternating field, ions of a given mobility arrive at C when the potential there is ready to pull them through on to the recording plate E.

Suppose that ions of only one mobility are present, then, if the current received at E is plotted against the frequency of the alternating potential applied at AB and CD a curve similar to that shown in Fig. 1.3b is obtained. This consists in effect of a series of several orders of diminishing intensity, for not only do ions get through when the alternating period is such that the ions travel from B to C exactly in one period, but also



when the frequency is twice, three times, etc., this fundamental frequency. If different ions are present each produces its own characteristic peaks. From the frequency at which each peak occurs the time taken to travel the distance BC under a known field can be calculated and so the value for the mobility can be derived.

The following considerations show how it is that high



accuracy can be achieved. Suppose that the alternating pulse is of the form shown in Fig. 1.4a. At the beginning of a pulse, the ions situated at A (Fig. 1.4b) move across AB and at the end of the pulse have reached F. Only those ions which pass B, i.e. those between BF, move across the main field,

since the pulse alternation keeps back the rest. The mobility and pulse frequency are such that the ions just reach C at the beginning of the next pulse, enabling the entire ion flash BF to pass through to the collecting plate E. At frequencies other than integral multiples of this, no ions will get through, and by making the pulse short, the width of the layer BF can be made small compared with AB. This results in a high resolving power and also suppresses the higher order frequencies.

The nature of the ions

The earliest observations, although only approximate. revealed a difference between the mobilities of the negative and positive ions in a given gas, the former being larger than the latter. From the very simple kinetic considerations which give the ionic velocity as v=Xet/2m, it follows that vp/Xshould be constant, where p is the gas pressure, since the mean time t between collisions is inversely proportional to the pressure. Early workers found this to be true over quite a wide range of pressure. This being so, it seemed that the approximate formula was obeyed and could therefore be employed for the determination of the mass of the ion. all cases the values found for the ionic velocities were such that the masses derived were greater than the molecular weight of the gas through which the ion was moving. The ion was therefore considered to consist of a cluster of gas molecules instead of being regarded as a single ionised molecule.

Since at the time the earlier work was carried out little was known about atomic binding forces, it was taken for granted that if, for example, ions were being produced in nitrogen, the original ionised nitrogen molecule would attract to itself neutral nitrogen molecules thus constituting a cluster with mass exceeding the molecular mass of nitrogen. A great deal of experimental work was carried out in connection with this point. The recent work of Tyndall has proved that much of the earlier results have very little meaning because of lack of attention to very strict conditions of purity. He proved that clusters form when impurity molecules are present, the cluster consisting of the original ion to which is attached a small group of impurity molecules. When the gases employed are very highly purified, and mobility measurements are made at

low pressures, the mobilities are found to be greatly increased, particularly for the negative ions, owing to the tendency to shed the cluster and, in fact, under such conditions the negative ions remain as free electrons.

The marked effect of impurity is shown in Fig. 1.5 in which the velocity is plotted against X/p for ions in air. This should lead to a straight line. With impurity present this is the case, showing that, as the pressure is reduced, the ion retains its clustered mass formation, but when the gas is very dry

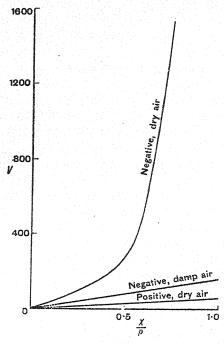


Fig. 1.5

and pure, the negative ion mobility increases at such a rate that the mass must be considered to diminish considerably as the pressure is reduced.

Each particular impurity introduced has its own specific effect. The mobility of negative ions is considerably reduced by the addition of polar molecules, i.e. molecules like H₂O, CH₃OH, C₂H₅OH, etc., the reduction increasing with the concentration of the impurity. When the aliphatic alcohols (methyl, ethyl, propyl, amyl) are used as the impurities

the amount of decrease in mobility is found to be proportional to the length of the molecular chain. This is to be expected if the clusters do actually consist of the impurity molecules, as the greater the molecular chain length, the greater will be the radius of the cluster and so the smaller the velocity (the effective mean free path is reduced). This view is strongly supported by the fact that non-polar molecules (e.g. decane), which can have little tendency to cluster formation, have very little effect on mobility. It is considered probable that each cluster is in a dynamic state, continually losing and capturing molecules.

The effect of age

Careful studies made in 1929 by Erikson upon the properties of positive ions a very short time after their creation have thrown considerable light upon the mechanism of cluster formation. In these experiments, the ions, immediately on formation, are dragged across an air stream by an electric field, the displacement down the stream during the time of crossing being an inverse measure of the mobility. By using high streaming velocities, mobilities could be measured when the age of the ion was as small as one five-hundredth of a second. With such short-lived ions it was found that both the positive and negative molecular ions have initially the same mobility. Any differences arising after ageing must be attributed to differential polar clustering. A progressive decrease in mobility takes place up to an age of about two seconds if even small quantities of polar impurity are present (water vapour is usually present in small amounts unless special precautions to eliminate it liave been taken). Since at normal pressure a single molecule can make as many as 1010 collisions during the 2 seconds, only a minute amount of impurity need be present for cluster formation to set in.

If the conditions of purity can be maintained sufficiently strictly to avoid the formation of clusters, then each foreign ion in a given gas should have its own specific mobility. If, for example, helium ions, or hydrogen ions, are sent through air, the mobilities measured should be characteristic for helium and hydrogen respectively in an atmosphere of air. The gas used must be very pure before this condition is arrived at, and as a further precaution observations must be made upon ions

whose ages are less than 2×10^{-3} second, thereby reducing practically to zero the chances of cluster formation. Under such conditions Tyndall finds that vp/X is constant over a very wide range, as the simple theory requires. Because of this it is possible to reduce all observed mobilities, determined under strict conditions, to a standard pressure (usually atmospheric) for the purposes of comparison. This has led to important conclusions.

The effect of ionic mass on mobility

The discussion immediately following concerns the mobilities of ions before clusters are formed, that is, it concerns specific mobilities of different ions. Tyndall and co-workers have made extensive observations on the mobilities of many types of ion in different gases. Table I exhibits a selection made from these results for ions of the alkali metals. The mobilities, in cm.²/sec. volt., were measured in helium, neon, nitrogen, and water vapour respectively as carrier gases, and refer to a pressure of 760 mm. Hg and 18° C.

TABLE I

Tons	Gas:					
	He	Ne	N ₂	H ₂ O		
Li ⁺	25·8 24·2 22·9 21·4	11.8 8.7 7.2 6.5	4·2 3·0 2·7 2·4	0·73 0·72 0·71 0·70		

For any given gas the mobility falls regularly with ionic mass. Fig. 1.6 shows the general relation between ionic mass and mobility found by Tyndall from observations upon sixteen different ions passing through nitrogen as carrier gas. The observed masses extend from that of lithium to that of thallium. The mobilities fall upon a perfectly smooth curve proving that the mobility is determined only by the ionic mass.

Useful information can be obtained by means of the massmobility relation about the nature of certain types of ion. When, for instance, hydrogen is introduced into nitrogen the ions which are formed, when the mixture is excited, have

mobilities which fall upon the curves at mass points corresponding to NH3 showing that the ions consist of NH3.

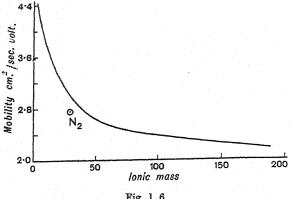


Fig. 1.6

The only deviation from the nitrogen mass-mobility curve is shown by nitrogen ions. There are theoretical reasons which explain why this should be so.

The relation between mass and mobility

A theoretical relation between mass and mobility was derived in 1905 by Langevin, based upon a simplified theory which assumed, for purposes of calculation, that the ions can polarise the gas molecules, which are regarded as elastic spheres. The mean free path of the ion in the gas can then be calculated and this leads to an expression for the mobility which is proportional to $(1+m/M)^{\frac{1}{2}}$ where M is the mass of an ion moving through a gas the weight of whose molecule is m. recognised that the theory is only approximate, yet in spite of this the observed variation of mobility with mass fits very closely into the Langevin formula. It is therefore justifiable to utilise this formula in discussing the question of clustering. Suppose we consider ions moving in an atmosphere of their own gas. When there is no clustering the ionic mass is equal to m, the mass of the gas molecule. Suppose next that large clusters form, so large in fact that the mass of the cluster M is very great compared with the molecular mass m (hence m/M can be neglected). The ratio of the mobilities in the two

cases is $(1+m/m)^{\frac{1}{2}}/(1+m/M)^{\frac{1}{2}}=\sqrt{2}$. From this it is seen that a very large increase in mass, due to cluster formation, has relatively only a small proportionate effect upon the mobility. Reference to Table I shows how large mass changes have only a small influence upon the numerical value of the mobility.

The effect of pressure and temperature

According to the simple kinetic theory the velocity of an ion is given by v=kX/p. The underlying assumption upon which this relationship is based is that the drift speed of the ion stream is small compared with the velocity of thermal agitation. As long as X/p is relatively small this is true, but as this increases a breakdown point is reached when the ionic velocity approaches that of thermal agitation. When this stage is reached it is necessary to use an empirical expression of the form $v=k(X/p)+k'(X/p)^2$ to give the drift velocity. The breakdown is usually critical, the smaller the ion mass the greater is the value of X/p at which the breakdown takes place.

Considerable experimental data on the effect of temperature have been reported but the general conclusions arrived at from the investigations have not been of great value. Tyndall has measured mobilities over a wide range of temperature $(80^{\circ}-500^{\circ} \text{ Abs.})$ and found, in accordance with theory that the mobility rises to a maximum and then falls. Empirically, the observed mobilities, k, can be fitted into an expression of the form $k=BT^{1}/(C+T)$ where B and C are constants and T the absolute temperature. This form of expression is similar to that derived by Sutherland for the effect of temperature

upon the viscosity of a gas.

The effect of polar impurities

By employing the relation $k \propto (1+m/M)^{\frac{1}{2}}$ the rate of growth of a cluster can be studied by measuring mobilities at different short intervals immediately after the birth of an ion. With a polar impurity present, say water vapour, it is found that the clusters grow very rapidly, the ion passing quickly into the stage of complete saturation. It is to be observed that the Langevin formula when applied to a cluster will only give an

"effective" mass, since there is an unknown change in the radius of the complex ion when the cluster has been formed. Strictly, the ionic radius is involved in the correct relationship. If the cluster is quite small and the pressure not too high the error introduced will not be considerable.

When the mass-mobility formula is applied to special cases the size of the cluster can be found. For example, Li⁺ ions in pure argon have a mobility equal to 4.99 cm.²/sec. volt., but the addition of I per cent. of water vapour as impurity reduces this to 2.26 cm.²/sec. volt. The increase in mass required to produce a reduction to this value corresponds to the attachment of six water molecules on to the original lithium ion. In general it is found that the number of molecules in a cluster is quite

small, varying perhaps from four to thirteen.

Special attention is to be drawn to the failures of the Langevin formula, the following being typical. The mobility of the pure nitrogen ion is 2.67 cm. 2 /sec. volt. in dry air, a value fitting into the mass-mobility curve. Nitrogen ion clusters which have six water molecules attached should have, according to the formula, the mobility 2.0 cm. 2 /sec. volt. This is the actual value found for clusters of short age when small traces of water vapour are present. However, at greater ages the mobility falls until it reaches the value 1.4 cm. 2 /sec. volt. It will be recalled that according to the formula a cluster of even infinite mass leads to a lower limit which is $2.67/\sqrt{2}=1.89$. The observed limiting mobility is much below this. The only explanation of the discrepancy that can be put forward is that the aged clusters are so large that the radius effect can no longer be neglected.

Ionisation by collision

From Figs. 1.1 and 1.2 it can be seen that increasing the field between a pair of parallel plates between which ions are being produced leads at first to saturation. With still greater fields saturation no longer occurs and the current increases rapidly until a state of instability sets in leading to a breakdown, with the passage of a spark, during which the resistance is low and the current high. The ionic process leading eventually to the production of a spark is called ionisation by collision. When an ion collides with a molecule or atom the collision can be considered to lead to a transfer of energy to the particle which

is struck. When the field is sufficiently high an ion acquires sufficient energy between collisions to enable it to ionise a gas molecule when impact takes place. The ion ejects an electron from the molecule it strikes creating thereby a pair of ions, one positive and the other negative. These newly formed ions are also rapidly accelerated by the field and in turn produce ionisation by collision, so that in a short time there is a very rapid geometric increase in the amount of ionisation between the plates (Fig. 1.7). The effectiveness of ionisation by collision

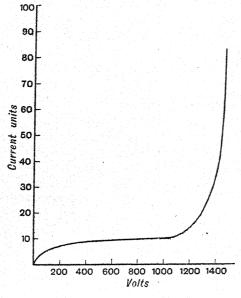
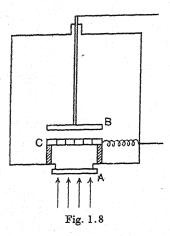


Fig. 1.7

will obviously be more marked at lower pressures since under these conditions the ions have time to acquire considerable energy before colliding with the molecule.

The detailed phenomena of ionisation by collision were first studied by Townsend with the apparatus outlined in Fig. 1.8. For accurate measurement it is necessary to produce ionisation in a plane and to do this use is made of the photoelectric effect. This will be more fully discussed later and for the present it need merely be mentioned that when ultra-violet light strikes

a metal surface, electrons are ejected from the metal. This, then, affords a convenient way for producing ionisation at a plane. In the apparatus shown in Fig. 1.8 ultra-violet light enters the experimental chamber through the quartz window A



and passes through the electrode C on to a flat zinc plate B which constitutes a cathode. C consists of a thin slab of quartz covered with a layer of silver, having transparent scratches upon it to allow the ultra-violet light to pass through. When the light strikes B, electrons are emitted, and as B is maintained at a negative potential, these are repelled and at once attach themselves to molecules, forming negative molecular ions. With the aid of guard rings a uniform field of several hundred

volts per centimetre is maintained between B and C. The apparatus contains a special mechanical device which enables B to be moved to and from C in a vacuum and yet remain parallel with C. A low specified gas pressure is maintained in the whole apparatus, and the current passing between B and C is measured for different field strengths.

Let the plates B and C be d cms. apart and let X be the

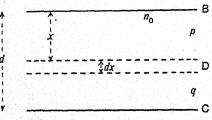


Fig. 1.8a

value of the electric field between them. The ultra-violet light generates a number of negative ions per second, n_0 , at the surface B. Suppose that by the process of collision each negative ion produces a new ions per centimetre of path.

Townsend assumes each positive ion creates β new ions per centimetre of path by collision. Consider a layer at D, dx cms. thick, and distant x cms. from B (see Fig. 1.8 α). Let p be the total number of pairs of ions generated per second between B

and D and q the number generated between D and C. In the element dx, the number of negative ions passing per second is the number originally produced at the plate surface, n_0 , together with the number generated between BD, i.e. p. The number of positive ions passing through dx in the opposite direction is only q. Hence the number of pairs of ions created per second in dx by collision is

$$dp = (n_0 + p)adx + q\beta dx \qquad . \qquad . \qquad . \qquad . \qquad (1)$$

If n is the total number of negative ions arriving every second at C we have,

substituting from (2) in (1) gives, on re-arranging,

$$dp/dx = (n_0 + p)(\alpha - \beta) + n\beta \quad . \qquad . \qquad . \qquad (3)$$

Integrating (3) gives,

thus

$$\log \left[(n_0 + p) + n\beta / (\alpha - \beta) \right] = (\alpha - \beta)x + \text{constant.}$$

This can be rewritten as

$$n_0+p=A\{exp(a-\beta)x\}-n\beta/(a-\beta).$$

The constant A can be obtained easily because p=0 when x=0, from which we get

$$A = n_0 + n\beta/(\alpha - \beta).$$

Also as $n=n_0+p$ when x=d, we find by direct substitution that

$$n = \frac{n_0(\alpha - \beta) \cdot e^{(\alpha - \beta)d}}{\alpha - \beta \cdot e^{(\alpha - \beta)d}} \qquad (4)$$

This formula fitted Townsend's observations and showed α to be much greater than β . Recent work has however shown that Townsend's proposed β mechanism of positive ion collision is quite incorrect. There seems no doubt now that β is not a measure of the power of the positive ions to ionise by collision. The β mechanism is highly complex. For instance it has been shown that each positive ion can create a number (γ) of new electrons at the cathode, by impact. On this basis an expression of the following form has been derived

$$n = n_0 \frac{(a - \beta')e^{ad}}{a - \beta'e^{ad}}$$

in which $\gamma = \beta'/\alpha$. Since β is in any event small compared with α , this cannot be distinguished experimentally from Townsend's equation. The actual γ mechanism is more complex than that indicated above, involving photo-electric processes also.

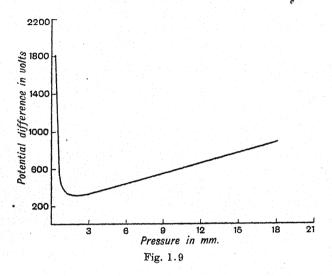
The spark discharge

It will be seen from the last formula that the number of ions passing between a pair of charged plates (and therefore the current too) would be infinite when $\alpha = \beta' \cdot e^{\alpha d}$, for then the denominator vanishes. Physically, this means that a spark passes. This takes place at a particular value of d, the distance between the plates, given by

Although this gives an approximately correct value for the sparking distance, the actual spark mechanism is much more complex than these simple formulæ suggest.

The obvious difference between a spark discharge and the earlier considered silent passage of current between two plates, is that the spark is explosive and luminous. The potential required to produce a spark depends upon the distance between the plates, the nature of the gas enclosed, and the pressure of the gas. Since the mechanism of ionisation by collision requires a definite finite time for the multiplication of a sufficient number of ions to carry the current of the spark. there is always a time lag between the application of a sparking potential and the actual breakdown. Further, it requires a greater voltage for the first spark to pass than for rapidly succeeding ones, since the first spark in its passage creates an abundance of ions in the vicinity of the electrodes. When the applied voltage only slightly exceeds the necessary sparking potential the observed time lag may be several minutes. If the spark gap is irradiated with ultra-violet light, photoelectric ions are generated and the lag is considerably diminished.

In making experiments upon the properties of sparks, a number of precautions require to be taken. Lag effects must be eliminated, and owing to the concentration of an electric field near a point, smooth plane parallel electrodes must be used. In the experiments carried out by Carr, plane parallel sheets of metal were embedded in an insulating material. The gas pressure between the plates could be varied, the potential needed for the passage of a spark being measured at different pressures. The curve found with a given plate separation is shown in Fig. 1.9. There is an optimum pressure,



depending upon the separation, at which a certain fixed minimum potential suffices to pass a spark. The shorter the distance between the plates the greater is the value of the critical pressure at which the minimum voltage appears. This minimum potential is characteristic of the gas used.

Paschen's law of sparking potentials

Paschen discovered that the sparking potential under any condition of pressure or plate separation depends only upon the product of the gas pressure and the distance between the electrodes. This is true for sparks passing at all pressures, and has been found to hold up to at least 50 atmospheres. The law can be derived in a general manner from collision

considerations. The value of the collision coefficient a is a measure of the tendency for an ion to produce ionisation by collision and clearly depends upon the velocity an ion acquires. To a first approximation the velocity of an ion at the end of the mean free path is Xet/m, and as t is proportional to 1/p where p is the pressure then a must be a function of Xe/pm. But a also depends upon the number of collisions which take place per centimetre and this is proportional to p, hence we may write

a=p. F(X/p)=p. F(V/pd).

F is some function and V is the potential difference between the plates. It can be shown, too, that independent of the β mechanism $\beta' = p \cdot F'(V/pd)$.

On substituting in (5), the formula for the sparking distance for a given potential, we get

$$pd = \frac{\log F(V/pd) - \log F'(V/pd)}{F(V/pd)} \quad . \quad . \quad (6)$$

From (6) it will be seen that the potential V required to produce a spark is a function only of pd. This is Paschen's law; it will be noted, also, that the mass of the gas between the electrodes is proportional to pd.

Brush and corona discharges

If either of a pair of electrodes maintained at a high potential difference is pointed, or has a radius of curvature at the end much less than the distance between the electrodes, a glow appears on the pointed electrode which changes into a brush-like form of discharge when the voltage used is a little below the sparking potential. This brush discharge can be seen even with currents as low as 2×10^{-9} amp. As the potential between the electrodes is slowly increased the brushes lengthen out into fine tentacles branching from each electrode. When the brushes from the two electrodes meet there is a sudden rush of current and a spark passes. If the potential is still maintained high, a steady burning arc keeps running between the electrodes. The direct cause of the brush is the concentration of electrical field in the vicinity of a point electrode. When this concentration reaches a given intensity, short sparks pass from the electrode and terminate in the air.

It is of interest to note that the ion current from a point at a high potential sets the air into violent motion, producing an "electric wind" which can easily be detected.

If one of the electrodes is a fine wire the brush takes on a form known as the "corona discharge." When a certain voltage is reached, either a continuous sheath or luminous beads surround the fine wire. This constitutes the corona, and when the voltage is increased beyond this point the corona breaks up into the ordinary brush form of discharge. Technically, the formation of the corona is of importance in connection with the losses from high voltage power transmission lines.

Condensation properties of ions

A striking property exhibited by all kinds of ions is that, like dust particles, they possess the power of precipitating condensation clouds from air supersaturated with vapours. When air which is saturated with water vapour is suddenly cooled by an adiabatic expansion, it becomes supersaturated, and if there are dust particles present, a condensation cloud forms. If the dust is removed by filtration no cloud is precipitated, from which it must be concluded that the dust particles act as nuclei. In dusty air, condensation droplets form on each particle of dust. The adiabatic expansion must be carried out rapidly if reproducible results are to be obtained, for a slow expansion allows the gas to warm up quickly and supersaturation disappears.

Wilson showed that moist dust-free air will also give a condensation cloud when expanded adiabatically to 1.38 times its original volume. Vapour pressure tables show that this is equivalent to creating an eightfold supersaturation of water vapour. When, however, ions are present in the dust-free air, a cloud appears at an expansion ratio of 1.25, which represents fourfold supersaturation only. It is clear that the individual ions must be acting as condensation nuclei, assisting the precipitation of the cloud. Wilson was further able to prove that the minimum expansion needed for cloud precipitation differs for positive and negative ions. For the negative ions an expansion ratio of 1.25 (fourfold supersaturation) suffices, but the positive ions require a 1.31 ratio (sixfold supersaturation). Unless the two kinds of ions have been

previously separated out by an electrical field, it is clear that the higher 1.31 ratio will bring down a cloud containing both positive and negative ions. The cloud formed on ions as nuclei will carry with it the charge of all the ions trapped in it.

Theory of condensation upon ions

It is easy to show that, to a first approximation, the vapour pressure of a spherical drop of liquid exceeds that of a plane surface by an amount $P=2T\sigma/r\rho$ where ρ and σ are the densities of the liquid and its vapour respectively, r the radius of the drop, and T the surface tension of the liquid. If r is very small then P will be very large, which means that if a very small condensation droplet forms, for some reason, the vapour pressure of the liquid in this drop will be so high that immediate re-evaporation will take place and the drop will disappear rapidly. Thus, even if there is a considerable amount of supersaturation, drops will only form if they find some nucleus to build upon which will enable them to start with a value of r sufficiently large to keep down the vapour pressure to a reasonable value. This explains why clouds form easily in unfiltered air and why an eightfold supersaturation can be reached when the air is dust free. The drizzle which sets in above this point takes place because the air molecules themselves behave as the necessary nuclei.

Consider now what happens if a condensed water droplet is given an electrical charge e. The electrical potential energy of such a charged droplet is $e^2/2r$ and, as this decreases when r grows bigger, the net effect of the charge is to make the drop expand in order to bring the electrical potential energy down to a minimum. On the other hand, the potential energy due to the surface tension of the liquid of the drop is $4\pi r^2 T$, which decreases when r diminishes. The surface tension thus tends to cause contraction of the droplet. 'It is clear, therefore, that the electrical charge acts in an opposite sense to the surface tension and thus brings about an effect which is equivalent to a reduction in surface tension of the liquid constituting the droplet. Since the excess vapour pressure isproportional to the surface tension, a charged drop will not evaporate so rapidly as an uncharged drop with the same radius, therefore an ion will enable cloud condensation to set in on itself before a normal uncharged molecule will do so.

The preferential action of the negative ions has its origin in the fact that water drops are known to possess a surface double layer of electrification with the negative coating inside. This makes negative ions more efficient than positive ions as condensation nuclei.

The Wilson cloud expansion chamber

The cloud expansion chamber devised by C. T. R. Wilson is an instrument of great simplicity and of great importance. It is shown in outline in Fig. 1.10. The cylindrical glass

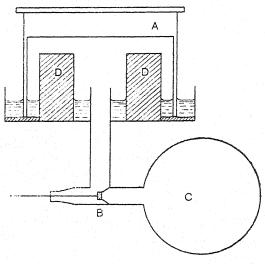


Fig. 1.10

chamber A stands upon a rubber disc in a trough which contains water. A movable piston produces the expansion in A. The blocks of wood D are introduced in order to reduce the air volume within the hollow piston. A large evacuated flask C is connected to the piston through a tube by way of the valve B. When B is opened the air under the piston is connected to the vacuum chamber C and as a result the piston drops suddenly, producing thereby an adiabatic expansion in the cylinder A. The expansion ratio can be adjusted by altering the height of the piston. In practice, several expansions are repeated until all the dust particles are precipitated

out of the air. If ions are then formed in A, an expansion will cause condensation upon the individual ions (see Plates IV and V).

IV and V).

Many ingenious devices have been used for improving the cloud chamber. The drops around the ions are now usually observed by photographic methods. Two cameras can be employed, set at an angle to each other, and from the combination of the two simultaneously taken pictures the position in space of the cloud can be determined. This is of great importance when the ion tracks produced by individual rapidly moving ionising particles are to be studied. Shimizu has designed a simple form of expansion chamber in which the expansion piston is operated by the rocking arm of a motor. The expansions are not so rapid as in the apparatus previously described, but by this means as many as three expansions per second can be made. A cinematograph camera is used in conjunction with this chamber resulting thus in a very great saving in time. A notable advance in the technique of the expansion chamber is due to Blackett, who has constructed an automatic chamber which is set into action by the passage of a single ionising particle. Thus the particle whose track is required acts as a trigger and makes the expansion take place. The result is that almost every expansion photograph contains tracks, whereas previously, in many photograph contains tracks, whereas previously, in many experiments, there was only a small chance of obtaining the sought for tracks when a given expansion was made. Fuller details of this chamber will be given later.

details of this chamber will be given later.

In Plate IVA a photograph is shown exhibiting the ion cloud tracks obtained when a beam of X-rays is passed through the cloud chamber. It can be clearly seen that along the path of the X-ray beam electrons are ejected from atoms. Each high speed ejected electron produces ionisation along its path, which turns and twists because of frequent atomic collisions. The path of each electron is made visible by the condensation of droplets upon the ions formed by collision with atoms. Plate IVC illustrates the tracks formed when individual very fast ionising particles are passed through the individual very fast ionising particles are passed through the expansion chamber. The separate droplets which form round each ion can often be counted. It is important to notice that in using photographic methods the exposure must be made a

THE CONDUCTION OF ELECTRICITY THROUGH GASES small fraction of a second after the expansion, otherwise the drops fall and the tracks become distorted and ultimately are destroyed.

REFERENCES

[&]quot;The Conduction of Electricity through Gases." J. J. Thomson and G. P. Thomson. Vol. I (1928), Vol. II (1933).

"Mobility of Positive Ions in Gases." A. M. Tyndall. (1938).

CHAPTER 2

THE ELECTRIC DISCHARGE THROUGH GASES AT LOW PRESSURES

Appearance at reduced pressures

When an electric discharge is passed between two electrodes in a tube which contains a gas at a low pressure, the appearance differs very much from that of the spark in air at atmospheric pressure. On applying the electric field and then steadily reducing the Pressure, the original crackling spark broadens out into a quiet streamer discharge. At a pressure of about 1 mm. of mercury most of the tube is filled with a

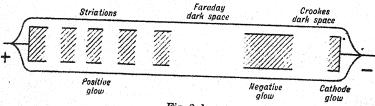


Fig. 2.1

glow that extends from the anode. This glow, called the positive glow, does not fill the whole of the tube but ends near the cathode, a dark space intervening. This dark space, the Faraday dark space, extends close up to the cathode which itself is covered by a thin glowing layer, the negative glow. On further reducing the pressure, the negative glow is seen to detach itself from the cathode and a second dark space, the Crookes dark space, appears between the negative glow and the cathode. Under these conditions the discharge has the appearance shown in Fig. 2.1.

When the pressure falls to about 0.1 mm. of mercury the glowing positive column splits up into alternate dark and light bands called striations, which become longer and fewer as the pressure is still further reduced.

When the gas pressure is reduced still further the Crookes dark space spreads out until it fills the whole tube and the glow in the gas itself disappears. At this point the electrical resistance of the tube increases considerably and it is difficult to pass any current at all if the pressure is still further reduced. Although there is no glow in the gas within the tube, the cathode is giving off a stream of charged particles, the cathode rays, which on striking the glass walls of the tube excite it into fluorescence. Sir J. J. Thomson first proved that this cathode emission consists of a high-speed stream of electrified particles—electrons—moving in straight lines from the cathode. This beam throws sharply defined shadows of solids interposed in its path, can exert a small mechanical pressure, heat bodies to incandescence, and excite fluorescence. The particles constituting the beam can be easily deflected by small magnetic fields and by electric fields, and carry a negative charge. When the beam is allowed to strike solid bodies the latter emit X-rays so that a discharge tube at a pressure sufficiently low to give a defined cathode ray beam always emits X-rays.

Potential variations in the discharge tube

The complex phenomena taking place in a low pressure discharge tube can be investigated by examining the electrical potential distribution across the tube. Many experimental methods have been devised for this purpose, that due to J. J. Thomson being illustrated in Fig. 2.2. The discharge to be studied passes between the electrodes A and B which constitute anode and cathode respectively. These are fixed on to an insulating base mounted so that the complete electrode system can be moved from one end of the tube to the other without altering the pressure. EF is a small subsidiary discharge tube which sends a fine pencil of cathode rays down between A and B on to the fluorescent screen S. This pencil of rays is subjected to the electrical field between A and B and is thus deflected in its passage through the discharge, the amount of the deflection depending on the strength of the field through which it passes. By moving successive portions of AB under the cathode ray pencil the whole field distribution in the discharge can be studied.

As is to be expected from the variation in the appearance of the discharge with pressure, the field distribution is very markedly affected by pressure changes. Two cases are

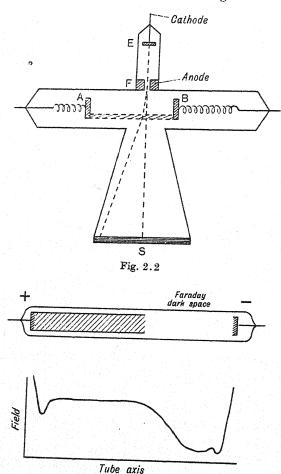
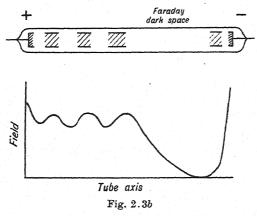


Fig. 2.3a

illustrated. In Fig. 2.3a the pressure is not sufficiently low to produce striations. The curve below the discharge tube shows the field distribution. There is a very steep fall of potential across the dark space close to the cathode. Most of the fall of potential across the tube as a whole takes place in

the dark space. In the discharge shown in Fig. 2.3b striations are in evidence and the main potential fall is across the cathode or Crookes dark space. The minor alternations which are observable correspond with the striations, each striation dark space showing a fall in the field.

Up to a certain critical limit of current, depending upon the size of the cathode, the cathode fall of potential is independent of the current and pressure in the discharge tube. In this condition the potential drop is called a *normal cathode fall*, its value being fixed largely by the nature of the gas used although it is somewhat affected by the electrode material. The normal



cathode drop exists only as long as the cathode remains partially covered with glow. When the current in a tube is increased, the area of glow on the cathode spreads so as to maintain a constant current density at the cathode. If the current is increased after the cathode is completely covered with the glow, a change in the potential fall takes place, the value increasing with the current. The potential fall is then described as anomalous. Normal cathode falls are of the order of several hundred volts.

The mechanism of the discharge

In order to understand the general mechanism of the low pressure discharge we shall consider the complex condition described in Fig. 2.3b. The whole appearance can be attributed to the occurrence of ionisation by collision within the tube. It is

clear that if ionisation of the gas molecules is produced by any means whatsoever, both positive and negative ions will be created in the process. At the instant of creation the negative ions will be electrons and the positive ions the residual positively charged molecules. Owing to the relatively high mobility of electrons, the latter will be repelled very rapidly from the neighbourhood of the cathode. There will thus be an accumulation of positive ions near the cathode, a fact which explains why the gradient of field here is intense and the cathode potential fall large. The accelerated positive ions strike the cathode and cause it to liberate a copious stream of Some of the moving ions strike gas molecules and ionise by collision, creating both types of ions. Owing to the large gradient produced by the accumulation of positive ions near the cathode, the electrons created in this neighbourhood are rapidly accelerated. Moving electrons give off no light at all, since there is only light emission from atoms or molecules which have been excited. As soon as the rapidly moving electrons strike molecules they ionise by collision and also cause excitation glow. It is clear, therefore, that the length of the Crookes dark space corresponds to the mean free path of the electrons, that is to say, they must on the average travel this distance before striking a molecule. The beginning of the negative glow should thus represent the start of ionisation by collision by the electrons. This is only approximately correct, for it seems probable that the dark space does in fact correspond to more than one mean free path.

However, the accumulation of the positive ions near the cathode, which originally assisted the electrons away from this region, now tends to reduce the speeds of the electrons which have passed through the dark space. These are soon slowed down until they are unable to ionise by collision. This is the beginning of the Faraday dark space. Quickly leaving the influence of the positive ions behind them, the electrons in the Faraday dark space now begin to accelerate towards the anode and again start to ionise by collision with the corresponding appearance of glow in the gas. This constitutes the beginning of the positive column, both positive and negative ions being of course produced there by the collision process. The electrons created at the beginning of the positive column have high mobility and move rapidly to the anode leaving behind a positive ion accumulation once more which slows down the

electrons until they are unable to ionise and thus once more a dark space is created. In this manner a striation is produced in the positive column. Hence the luminous portions of the striations correspond to the negative glow and the dark regions to the Faraday dark space. It is clear that the striation formation process can repeat itself until the anode is reached. The striations must lengthen out with reduction in pressure since the mean free path then increases.

The above description is not meant to be complete since it does not account for all the detailed phenomena that can be observed in discharge tubes. Many different types of striations have been observed, some sharp edged, some multiple, some moving, some stationary, etc. The shape of the discharge tube and of the electrodes has a definite effect upon striation formation. The striations are very easily excited if the gas in the tube contains carbon compounds such as CO, CO_2 , or organic compounds like C_6H_6 , etc.

Cathode rays

When the pressure is so low that the dark space extends and fills the whole tube, patches of greenish fluorescence appear on the glass opposite the cathode. These patches are caused by a radiation emitted from the cathode surface. If a solid body is interposed in the path of the rays, between the cathode and the tube walls, a sharp shadow is cast in the fluorescent light patch. From the size and sharpness of the shadow it is evident that the radiation travels in straight lines and is emitted normally from the cathode surface. The beam of radiation is called a cathode ray beam, and it can be proved that it consists of a stream of rapidly moving electrified particles—electrons.

A very large number of bodies phosphoresce when cathode rays are allowed to fall upon them and by making use of this fact the rays can easily be detected. In particular, crystals of potassium platino-cyanide or willemite phosphoresce very brightly, and if a cathode ray pencil is projected on to a screen covered with a fine powder of this material, a bright spot of light is formed. The movements of this spot can be observed if the cathode ray beam is subjected to electric or magnetic forces and by this means the properties of the rays can be studied. Since

the rays are emitted normally to the cathode surface, they will be brought to a focus if the cathode is shaped like a concave mirror. The rays are able to heat bodies by impact and if brought to a focus can even melt platinum. A secondary effect associated with the heating is that cathode rays can produce mechanical rotation of a suspended vane held in their path. This is due to a radiometer action, one side of each vane being heated more than the other and thus causing bigger reaction from the more active rebounds of gas molecules coming into contact with the hotter face. Apart from this, there are indications that the cathode rays do exert a small true mechanical pressure which can be detected if the gas pressure is very low, the radiometer effect then being eliminated.

When a bar magnet is brought up to a discharge tube in which a cathode ray beam is falling upon a fluorescent screen, the spot of light moves in one direction or the other according to which magnetic pole is presented to the rays. It is found that quite small magnetic fields suffice to deflect a beam of the cathode rays. If the rays are passed between a pair of metallic plates maintained at a potential difference the luminous spot is also deflected. The deflection by means of magnetic or electric fields is in accordance with what would be expected of a stream of charged particles.

An important experiment carried out by Perrin in 1895 shows that the rays carry with them a negative charge. The form of the apparatus used is illustrated in Fig. 2.4. The cathode rays are produced in the side tube A, a fine beam being selected by means of the aperture B. The cathode ray pencil causes a fluorescent green spot on the wall of the large bulb C. On bringing up a magnet the patch of light C can be slowly deflected towards the opening of a metal cylinder D which is surrounded by an earthed shield. D is connected to an electrometer, and as soon as the cathode rays are seen to fall upon the opening, a large negative charge is registered, proving that the cathode rays carry a negative charge.

Hertz first noticed that cathode rays can penetrate thin metal foils. Lenard studied this property more thoroughly. A valuable property of cathode rays is that they have the power of blackening a photographic plate, like light, thus

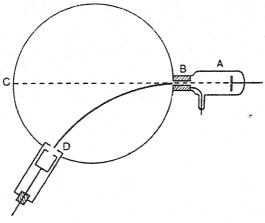


Fig. 2.4

deflections of cathode ray beams can be studied photographically and a high degree of accuracy can be achieved.

Cathode sputtering

.When a discharge tube has been run for a considerable time the glass near the cathode becomes covered with a deposit of metal often in the form of a mirror. The cathode metal is said to have "sputtered." Cathode sputtering is technically an important method for the production of thin metal films and high quality mirrors. The discharge tube can be designed so that the sputtering from, say, a silver cathode is quite uniform and by this means optically perfect uniform thin films can be made. The mechanism of sputtering is highly complex and is not yet fully understood. It is, however, quite certain that the phenomenon arises from the disintegration of the metal surface of the cathode due to the bombardment by the positive ions attracted to it. It is considered that the sputtered particles consist largely of uncharged metal atoms travelling with velocities corresponding to the thermal velocities of the atoms of the vapour of the metal at its boiling point.

The efficiency of the sputtering action depends upon the molecular weight of the gas used in the discharge tube. The heavier the positive ions the more effective is the dis-

ruptive action on collision with the cathode surface. A regular increase is observed in the amount of metal sputtered in a given time when the molecular weight of the gas in the tube is increased. Argon, for instance, is sixteen times more effective than hydrogen. Some metals, amongst which are, for example, silver, gold, cadmium, etc., sputter with great ease, but others, like aluminium, magnesium, etc., hardly sputter at all. The failure to sputter is considered to be due to the existence of a protective oxide coating upon the non-sputtering metals.

Sputtering in discharge tubes can become a source of trouble since the walls coat with metal and become opaque to light. As the thin metal layer is electrically conducting sparks sometimes jump to the glass walls and either short circuit the discharge or even cause a puncture in the glass wall.

Types of low pressure discharge

"Vacuum" discharge tubes of diverse forms are of great importance in many branches of science and industry. To name but a few applications, for instance, there are (1) X-ray tubes, (2) spectroscopic discharge tubes, (3) "vacuum" rectifiers, (4) oscillographs, (5) neon advertisement lighting, etc. The need for so many different tubes has resulted in the growth of a special branch of electrotechnics which concerns itself with "vacuum" tube discharges. We shall consider here a few of the specialised types of discharge which have been developed for the purely scientific purpose of studying spectra, since the investigation of the latter has led to important knowledge concerning the structures of atoms and molecules.

The Geissler tube

The Geissler tube, which has been used for many years as a spectroscopic source, can take on many forms. A simple kind is that shown in Fig. 2.5a. The tube is in the form of a letter H, having a capillary portion AB about 4 cms. long, with a bore of the order of 1 mm. The electrodes CD are often of platinum. The tube is filled with the gas whose spectrum is required, the pressure being 1 or 2 mm. of mercury. An alternating potential exceeding 2,000 volts is applied either from a spark coil or a transformer and the passage of the discharge results in a concentration of the positive column in

the narrow tube AB, the current density being high here.

When the capillary tube is viewed "end on" a very bright emission can be observed. Such tubes have proved of great value in connection with the study of the line spectra given out by atoms and the band spectra given by molecules. Owing to the fact that fairly large electrical potentials are required, the lines of the spectrum are all somewhat broadened, since strong electrical fields affect the motions of electrons in atoms and molecules. This in-

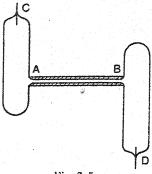
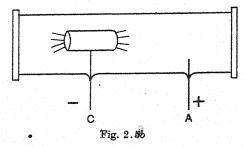


Fig. 2.5a

troduces difficulties in certain investigations requiring lines which are extremely narrow. When extremely sharp and narrow lines are required, the source used is a "hollow cathode" discharge.

The hollow cathode discharge

The hollow cathode discharge, discovered by Paschen, has led to very important information concerning the nuclei of atoms. It is a means for producing spectral lines with great intensity and of extreme sharpness. The simplest form of the tube is shown in Fig. 2.5b. The anode A can consist of a



piece of platinum or nickel wire, the cathode C is in the form of a hollow metal cylinder. The tube is operated with direct current, a potential exceeding 500 volts (preferably about 1,500 volts) being required. The gas pressure in the tube is critical, being of the order of 1 mm. of mercury or so, the actual value depending upon the dimensions of the cathode.

When the pressure is correct, the application of a direct voltage supply between A and C produces a peculiar discharge. The inside of the hollow cathode fills with a brilliant glow, particularly if a large current is passed. The spectrum of the hollow cathode glow consists of both the spectrum of the gas which carries the discharge and of the metal of the cathode. By introducing materials within the hollow cathode, their spectra can be excited and studied.

The mechanism of excitation is very complex and appears to be largely due to a sputtering action. The positive ions in the gas discharge bombard the cathode, or the material on it, and this excites the spectra. Since the light is virtually excited within a Faraday cylinder, that is within a place where the electrical field must be very small, the emitted lines are very sharp. This description is however but crudely correct. This remains true as long as there is a normal cathode drop, which exists as long as the current is such that the whole of the cathode is not completely covered with glow. The cathode can be cooled, and under these conditions of low temperature and absence of strong field, the emitted lines are very narrow. The study of the inherent close hyperfine structure of these narrow lines has led to important discoveries about atomic nuclei.

The electrodeless discharge

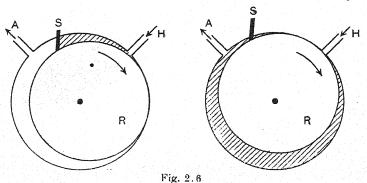
There are two distinct types of "electrodeless" discharge, these being called the "ring" discharge and the "high frequency" discharge, respectively, although both are actually high frequency discharges. In exciting the ring discharge, which was discovered by J. J. Thomson, a partially evacuated tube which possesses no electrodes at all is placed inside a solenoid through which is passed a high frequency alternating current, produced by the oscillatory surge created when a Leyden jar is discharged through the coil. The lines of electric force form closed curves, so in this case the discharge takes the form of a bright ring of light, coaxial with the solenoid. It is very energetic, especially if a quenched spark is in series with the coil and Leyden jar. Owing to the violence of the discharge, many electrons can be torn from atoms, producing thus a state of multiple ionisation. Multiply ionised atoms can readily be studied by this means.

The action of the "high frequency discharge" is very different from that of the ring discharge. It is excited by means of the electromagnetic waves given out by a valve oscillator, radio frequencies being required. Wavelengths between 1 and 10,000 metres are equally capable of exciting the discharge (frequencies 3×10^8 to 3×10^4 per second). A "vacuum" tube without electrodes is held close to the oscillating circuit, or strips of foil wrapped round the outside of the tube are connected to the inductance coil of the oscillator. The tube glows brilliantly, and although the discharge is very intense, the excitation is quite mild, being effectively a low voltage excitation. The particularly interesting property exhibited by this form of discharge is that it will pass easily through a tube in which the pressure is so low that an ordinary electric discharge cannot pass at all. Thus it is possible to study the behaviour of atoms at very low pressures where collisions are infrequent.

The production of high vacua

Most of the advances in modern physics have been due indirectly to the ease with which high vacua can be produced. Mechanical rotary pumps of the inexpensive type can rapidly produce a vacuum with a pressure of 0.001 mm. of mercury. This suffices for many experimental studies. Rotary molecular pumps can attain much lower pressures and are of very great value, but for general purposes a rotary oil pump is used in conjunction with a diffusion pump.

The mode of operation of a common type of rotary pump is illustrated in Fig. 2.6. The construction is entirely of



metal. An eccentrically mounted rotor R sweeps air from the side to be evacuated H, to the atmospheric side A. At the beginning of each stage of pumping air from H enters the shaded chamber and as the rotor moves round this trapped air is swept round to the opening A. Further completion of the rotation stroke compresses the air and drives it out into the atmosphere, S is an oiled spring valve dividing the high and low pressure sides. Frequently pumps have two rotors, out

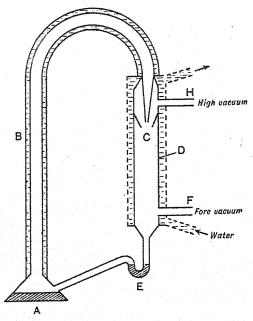


Fig. 2.7

of phase, mounted on the same spindle and connected in series, one pumping behind the other in the form of a double stage pump. The whole rotary system is immersed in low vapour pressure oil. With a speed of rotation of about 250 revolutions per minute a vacuum of 10^{-3} mm. of mercury can be attained.

In order to reach pressures of 10^{-6} mm. and lower, and such vacua are frequently required, diffusion pumps are generally used in modern practice. A simple type is shown in Fig. 2.7. Mercury or low vapour pressure oil contained in A is heated and the

vapour, after passing up the lagged tube B, shoots out of a jet C and then condenses upon the walls of the water-cooled chamber D. The condensed mercury, or oil, runs back to A via the trap E. The apparatus to be evacuated is attached to the side limb H. A rotary oil pump is connected to F, since the diffusion pump will not operate without an auxiliary fore-vacuum pump.

The action of the pump is complex. The preliminary reduction in pressure produced by the rotary pump permits the vapour jet issuing from C to attain a very high velocity. Gas diffusing from the apparatus being exhausted is trapped by the vapour jet, carried down to the rotary pump, and removed. The velocity of the vapour jet is much greater than the rate of diffusion of gas molecules, hence there is virtually no chance of any back diffusion into the vacuum chamber. The speed of the pump, and the fore-vacuum required, depend upon the shape and size of the jet C. When high speeds and low pressures are required multiple stage pumps are employed, consisting of multiple tiers of jets in series, each acting as a fore-vacuum pump for the jet following.

The ultimate vacuum depends upon the vapour pressure of mercury or of the oil employed. With greased vacuum taps in the system, the tap grease sets a limit to the pressure that can be reached. There is always a constant slow "evaporation" of occluded gases from the walls of any vacuum system and this decides the final limit that can be attained, if greases are avoided, and even if the pressure of the oil or mercury vapour is kept extremely low by means of liquid air traps.

A convenient rapid alternative method for producing a high vacuum in a chamber of small dimensions is by the employment of charcoal cooled by liquid air. This process cannot be applied generally, but where it can be used it is extremely efficient. The pumping action is based upon the fact that specially prepared charcoal, when cooled to liquid air temperature, has an enormous capacity for absorbing most gases. The absorbing power depends upon the manner in which the charcoal is prepared, that made from the shells of coconuts being by far the best. The charcoal is activated by leaving it open to the air whilst it is cooled by liquid air. It absorbs many hundreds of times its own volume of air and the pores in the charcoal are loosened. The absorbed gases are then

pumped off, the charcoal being heated to at least 350° C. during the pumping. On cooling down it is in a state suitable for use as a pump. A tube containing the activated charcoal is attached to the vacuum system, which is then evacuated to a preliminary stage with a rotary oil pump. The latter is disconnected and on immersing the charcoal tube in liquid air, pumping at once begins. The charcoal absorbs most of the residual gases and a high vacuum results.

Charcoal is not able to absorb all gases equally well. Helium is hardly absorbed at all and neon very slightly. This differential action is of value, for by its means it is easy to isolate and purify samples of rare gases.

Extremely good vacua are frequently required in industry. The thermionic valves used in radio, for example, require to be highly evacuated if a good response is to be expected. To obtain these vacua the "getter" action of various metals is employed. The apparatus is first evacuated with a diffusion pump. A small quantity of calcium metal is introduced into the apparatus and after the glass parts are sealed off from the pumping system this metal is rapidly evaporated ("flashed" is the technical term), forming a mirror deposit on the walls. A very large effective surface is thereby exposed. By chemical surface adsorption the last traces of gas are removed. original gas being air, the residual gases are almost entirely nitrogen, oxygen, water vapour, carbon dioxide, and hydrogen. These form calcium nitride, calcium oxide, calcium carbonate, and calcium hydride. (The name "getter" has arisen in industry since the metal "gets" the residual gases.)

It must be realised that the term "vacuum" is only a relative expression. Although a pressure of 10^{-6} mm. Hg is only 1.3×10^{-9} of an atmosphere there still remain 3.5×10^{10} gas molecules per cubic centimetre in such a vacuum. The physical significance of this low pressure is more readily appreciated if it is recognised that under such conditions the mean free path of the gas molecules may be several metres. This means that in small vessels beams of atoms, electrons, or ions can be considered to travel as if in free space.

REFERENCES

[&]quot;The Conduction of Electricity through Gases." J. J. Thomson and G. P. Thomson. Vol. I (1928), Vol. II (1933).

CHAPTER 3

THE CHARGE AND MASS OF THE ELECTRON

Introductory

The studies of J. J. Thomson on the electrical discharges in gases first proved that the cathode rays consist of streams of electrified particles each some two thousand times lighter than the atom of hydrogen. The particles, electrons, are fundamental constituents of every type of atom. It will be assumed for the moment that the mechanical and magnetic properties of the cathode ray beams constitute sufficient proof that they consist of streams of charged particles. In the experiments now to be described it is shown that all the cathode ray particles are similar in nature, having the same charge and mass but possibly having different velocities.

As electrons are sub-unit components common to all atoms the exact determination of the electronic charge and mass is of great importance. These quantities are fundamental atomic constants and theory shows that they appear in very many formulæ for physical relationships. The first approximately accurate measurements were made in 1897 by Thomson, and since then successive refinements have led to greatly improved values for both the charge and the mass. Many entirely different high precision experimental methods have been devised and all lead to closely similar numerical results (these will be discussed later in an appendix). The earlier historically important experiments will only be briefly described since they have now been superseded by much more accurate determinations.

The cloud chamber determination of the electronic charge (Thomson)

The cloud chamber employed in this experiment is that designed by C. T. R. Wilson. Air saturated with water vapour

is brought into a state of supersaturation by a rapid adiabatic expansion. Wilson had shown that if there are ions in the air space of the chamber, a single water drop condenses upon each of these ions when the expansion ratio is suitable. In Thomson's experiments a condensation cloud was precipitated upon negatively charged ions by adjusting the amount of expansion. The cloud was observed to fall slowly under gravity and the rate could be measured. It is known that a drop falling in a viscous medium (in this case the air) obeys the Stokes' formula, according to which a drop of radius a and density ρ attains a constant terminal velocity given by $v=2g\rho a^2/9\eta$ where η is the coefficient of viscosity of air. From observations upon the rate of fall of the top of the expansion cloud the mean radius of the drops can be calculated.

Since the expansion ratio employed to produce the supersaturation is known, the change in temperature due to this expansion can be simply derived from the gas laws. With the aid of vapour pressure tables the amount of water vapour that has gone over into drops can therefore be calculated. As the radius of each drop is known the total number of drops in the cloud is obtained. The cloud is allowed to settle on to a surface connected to an electrometer which measures the total charge carried by the drops. Thus the charge carried by each drop, which is the charge on a single negative ion, is arrived at.

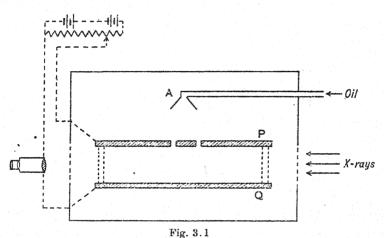
Although of historical importance, the above method may now be considered obsolete since there are many inherent inaccuracies associated with it. Thus the gas warms up rapidly after the expansion and the drops evaporate, turbulence effects exist, the drops have not all exactly the same size, the rate of fall of the diffuse cloud edge is difficult to measure, etc. We shall therefore consider now a precision method for measuring the electronic charge, due to Millikan.

Millikan's determination of the electronic charge

The method adopted by Millikan is an improved modification of an earlier experiment carried out by H. A. Wilson. In order to eliminate any errors due to variable drop size and to evaporation, observations are made on a *single drop* of low vapour pressure oil. A schematic outline of the experimental

20918

arrangement is shown in Fig. 3.1. The essential observation chamber consists of a pair of optically worked flat circular plates of metal, set accurately parallel with optically worked glass separators. Small holes are drilled into the upper plate, close to the centre and these enable oil drops from the spray atomizer A to enter the space between the plates. The drops, which are charged by friction during the spraying process, fall under the action of gravity with a velocity $v=2g\rho a^2/9\eta$. This can be measured by means of a microscope which has a scale in the eyepiece. If now a potential of about 1,000 volts is applied between the plates PQ, the upper plate being made



positive, the fall of a negatively charged drop will be retarded by the field. By adjusting the field strength a given drop can be brought to rest and maintained stationary in the field of view of the observing microscope.

In general a drop will have acquired a number of units of charge, each being equal to e, the electronic charge. Suppose that on removal of the electric field the drop under observation falls with velocity v_1 . Let a retarding field X be applied such that the drop begins to move upwards with a velocity v_2 . If the drop has n units of charge the net upward force is Xne-mg where m is the mass of the drop. It follows, therefore, that

$$\begin{array}{ccc} (\mathrm{X}\mathit{ne}-\mathit{mg})/\mathit{mg}=\!v_2/v_1 \\ \mathrm{giving} & \mathit{ne}=\!\mathit{mg}(v_1+v_2)/\mathrm{X}v_1. \end{array}$$

By means of an X-ray tube, which can send ionising radiation between P and Q, additional ions can be produced between the two plates. When one of these collides with the oil drop the charge alters abruptly and the observed velocity v_2 changes suddenly by a fixed amount $\pm v_0$ at each collision. Since at each alteration the drop acquires a single positive or negative charge by colliding with the appropriate ion, n alters by ± 1 . From this it follows that $e=mgv_0/Xv_1$. The same value for v_0 was found from thousands of separate observations on a single drop, proving that the charges upon the positive and negative ions are identical, apart from the sign.

To obtain e it is necessary to measure m. This can be obtained from v_1 , the rate of fall in air with zero field, since this is the terminal velocity in the Stokes' law expression for the fall of a spherical droplet in a viscous medium. According to this law the force due to gravity $4\pi a^3 \rho g/3$ equals $6\pi \eta a v_1$. This leads to a value for e which is

$$e = (9\eta/2)^{\frac{3}{2}} (v_1/\rho g)^{\frac{1}{2}} 4\pi v_0/3X$$
.

All the quantities in this expression can be measured.

Millikan noticed that the value derived for e was apparently not constant, as expected, but appeared to be greater for smaller drops than for larger. He proved that this inconsistency arose from the incompleteness of the Stokes' formula which requires a correcting term depending upon the mean free path λ of the molecules of the gas through which the droplet is falling. This correction is important when very small droplets are employed. The final corrected expression for e is

$$e = (9\eta/2)^{\frac{3}{2}} (v_1/\rho g)^{\frac{1}{2}} \{1/(1+A\lambda/a)\}^{\frac{3}{2}} 4\pi v_0/3X.$$

In this the constant A is obtained from a graphical plot of the apparent variation of e with a. The value found by this method, when using the most recent value for the viscosity of air, is

$$e = 4.8036 \times 10^{-10} \text{ e.s.u.}$$

This experiment is of fundamental importance, not only because of the accurate numerical result but because it proves in a very decisive manner that electricity is atomic.

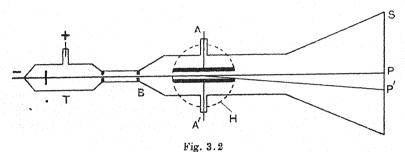
The many other indirect methods available for the determination of e are discussed in an appendix. The best mean value for the charge given by all these methods is

$$e=4.8025\times10^{-10}$$
 e.s.u.

The ratio of the charge to the mass of the electron (Thomson)

Thomson also first determined the mass of the electron. He devised an experiment for measuring e/m, the ratio of the electronic charge to the electronic mass m, and since the value of $e\cdot$ had already been obtained, m was deduced and found to be some two thousand times as small as that of the mass of the hydrogen atom. Since, as in the case of the determination of e Thomson's method has now been superseded, it will only be described very briefly.

The quantity e/m is determined by subjecting a narrow pencil of cathode rays to deflection by electric and magnetic fields which are at right angles to each other. The arrangement used is shown in Fig. 3.2. A fine cathode ray pencil



from a subsidiary tube T traverses fine slits and passes through a region in which there is an electric field X, produced by plates AA' maintained at a potential difference, and a magnetic field H produced by an electromagnet. The lines of force of the two fields act over the same region and are mutually perpendicular to each other and to the direction of the incident cathode ray beam.

Assuming that the cathode rays consist of charged particles with charge, mass, and velocity equal to e, m, v respectively, then the force exerted upon each electron by the magnetic field alone is Hev. This acts perpendicularly to the direction of motion and deflects the electrons into a circular path of radius R so that $Hev = mv^2/R$. The path is circular in the field only and if this extends over a distance l a particle moves through this path and then travels tangentially on leaving the field. If L is the distance from the centre of the field to the fluorescent

screen S, it can be shown from the geometry of the electron path that R=Ll/PP' where PP' is the observed (small) deflection of the spot of light due to the cathode pencil striking the screen. Hence

$$PP'=HelL/mv (1)$$

When the electric field alone is applied a force Xe acts in a vertical direction on the electrons producing an acceleration Xe/m acting during the time t spent in traversing the field. This time t is clearly equal to l/v where l is the length of the field and v the velocity of the electrons. The vertical velocity is therefore Xel/mv. On leaving the field the particles possess this vertical velocity and the horizontal velocity v, hence PP''/L is the ratio of these velocities where PP'' is the resulting (small) deflection, i.e.

$$PP''/L = Xel/mv^2$$
 (2)

Combining (1) and (2) gives

$$v = PP' . X/PP'' . H$$

 $e/m = X . (PP')^2/H^2 . PP'' . l L$

and

It is to be noted that if the field strengths are adjusted to give equal and opposite deflections (i.e. zero net deflection) then

$$Xe = Hev$$
 i.e. $v = X/H$.

It is not possible in practice to obtain perfectly uniform fields. When this is done the expressions for the deflections geometrical limits of the electrodes and pole faces. It is therefore necessary to integrate over the whole of the variable fields. When this is done the expression for the deflections are modified to

$$PP' = \frac{e}{mv} \int_{0}^{BP} \int_{0}^{x} (Hdx)dx$$

$$PP'' = \frac{e}{mv^{2}} \int_{0}^{BP} \int_{0}^{x} (Xdx)dx.$$

The integrals are evaluated by direct plotting of the field strengths along the length of the cathode ray beam.

Dunnington's method for e/m.

A number of high precision methods for the evaluation of e/m are described in brief in an appendix on fundamental constants. We shall deal here with one of these, the magnetic deflection method due to Dunnington, which is one of the most accurate methods yet devised.

The apparatus used is described schematically in Fig. 3.3. O is a high frequency oscillator supplying an alternating

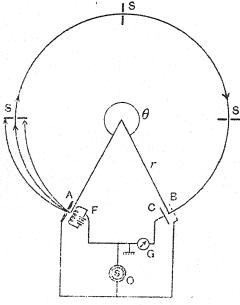


Fig. 3.3

potential at high frequency simultaneously to the electrode pairs AF and CB. F is a filament which emits electrons when heated, and at every half cycle these are accelerated towards A and pass through a hole in it. The whole apparatus is highly evacuated and maintained in a very large uniform magnetic field H. By means of a system of slits SSS placed accurately upon a circle, electrons with a specific velocity are selected from all those passing through A. After deflection through a circular path by the magnetic field, which has its lines of force perpendicular to the plane containing the centres

of the slits, the electrons are trapped by the Faraday cylinder C providing they succeed in passing through the gauze B, and are recorded by the galvanometer G.

If the fields are adjusted so that the time taken for an electron to pass through the angle θ exactly equals the period of the oscillator T (or is n times this), no electrons will succeed in reaching C, for at that instant B will have upon it the exact potential that A had when the electrons were accelerated by the latter. B therefore allows electrons to pass through (it is an open gauze) but draws them back before they reach C. If r is the radius of the circular path and v the velocity of the electrons

$$v = r\theta/nT = r\theta\nu/n$$

where ν is the frequency of the oscillator. We have also

$$v = Hre/m$$

as the particles are deflected by the field of value H. Equating gives

 $e/m = \theta \nu/nH$.

In carrying out this experiment H is varied until the galvanometer registers zero current, this value of H giving e/m. By using piezo electric quartz crystal control the oscillator frequency can be maintained constant to one part in a million. The angle θ can be measured with precision, so that the final accuracy of the determination depends only on the uniformity of the magnetic field and the evaluation of the field strength. By employing a very large powerful magnet Dunnington achieved an accuracy of one part in 4,500 finding

$$e/m = 1.7597 \times 10^7$$
 e.m.u.

At present ten precision determinations of e/m have been made, the weighted mean being

$$e/m = 1.7592 \pm 0.0005 \times 10^7$$
 e.m.u.

The mass of the electron

The same value of e/m is found for the electron no matter what source is employed for creating the electrons, provided that the velocities of the particles are small compared with the velocity of light. (It will be shown later that at high velocities the electron mass increases according to a law given by the

Theory of Relativity.) By combining the experimental values for e and e/m the value found for the electron mass is

 $m = 9.1066 \times 10^{-28}$ gm.

This will now be compared with the mass of the hydrogen atom. The latter is best found by determining Avogadro's number, which is the number of molecules in the gram molecule (or alternatively from Loschmidt's number, the number of molecules per cubic centimetre, since the weight of a cubic centimetre of hydrogen is known).

Determination of Avogadro's number (Brownian motion)

The first significant determination of this important constant was made by Perrin (1908) from a study of Brownian motion. The botanist Brown discovered (1827) that very small particles suspended in a liquid are in a constant state of agitation, due to thermal kinetic bombardment by liquid molecules. On the average, over a relatively long period, a particle receives an equal number of impacts in all directions, but if the particle is small, then over any given short time more particle impacts will in general be made in any one direction than in the directly opposite direction, as the impacts are governed by the laws of chance. The unbalanced impacts make the particle move in a path which is quite haphazard. Very high magnifications are needed to render the motion visible. This Brownian motion is perpetual and gives direct evidence for the truth of the kinetic theory of matter.

The suspended particles are in effect partaking of the thermal kinetic energy of the liquid molecules. When any heavy molecules are added as an impurity to light molecules (e.g. bromine added to hydrogen) then, in accordance with the equipartition theorem of kinetic theory, the thermal energy is equally divided between all the degrees of freedom in the mixture. In like manner suspended particles (either in a gas or a liquid) can be considered to be "molecules" of very high molecular weight, the weight of the particle being in fact the weight of the effective molecule. A suspension may therefore be regarded as equivalent to a low-pressure gas, since the number of particles per unit volume is small. The ordinary gas laws are applicable both to the distributions and motions

of the particles. It is possible to derive Avogadro's number N from a study either of the density distribution of the suspended particles or from their wanderings under the impacts of the liquid molecules.

Determination of N from density distribution of suspended particles

Under the combined action of gravity and the forces producing the Brownian motion, suspended particles in a liquid take up a vertical logarithmic density distribution similar to the density distribution of air molecules at varying heights in the atmosphere. This distribution can readily be derived by the following method.

Consider two horizontal planes with one a distance dh directly above the other. If ρ is the density of the gas (effectively constant over the small range) then the pressure difference between the two planes is

$$dP = \rho g dh = g dh M/V$$

where M is the gram molecular weight of the gas and V the volume of a gram molecule. Substituting from Boyle's law (PV=RT) gives dP=dh.MgP/RT.

Integrating between vertical points h cms. apart, at which the pressures are P_1 and P_2 respectively, gives

$$\log (P_1/P_2) = hMg/RT$$
.

Since the pressures are proportional to the number of particles per cubic centimetre N_1 and N_2 we have

$$\log (N_1/N_2) = hMg/RT.$$

This general expression applies to all "perfect" gases. For a gas, or suspension of particles each of mass m, the molecular weight is Nm where N is Avogadro's number. The effective mass of a particle of mass m and density ρ_1 suspended in a liquid of density ρ_2 is $m(\rho_1-\rho_2)/\rho_1$, hence

$$\log (N_1/N_2) = Nmgh(1 - \rho_2/\rho_1)/RT.$$

The number of particles per cubic centimetre is on the average proportional to the number passing through a plane of given area in a given fixed time (it will be remembered that the particles are in motion). If, therefore, the particles are viewed by a microscope set at the two points in succession, counts of the number of particles passing through the field of view in a

given period of time are proportional to N1 and N2.

Perrin employed a suspension of approximately spherical colloidal particles of gamboge yellow, of known density, and measured the radii by direct evaluation of the length of a chain of particles end to end. Uniformity in radius was obtained by means of preliminary centrifuging of the particles. By means of this experiment the first approximately correct value for N was obtained.

Determination of N from the motions of suspended particles

It was shown by Einstein that not only could the distribution of the particles yield a value for Avogadro's number, but in addition the latter could be derived by observing the linear motion of a particle during a measured time period. A single particle moves in a haphazard manner and after a given time t is displaced a distance x from its original position. The force opposing the motion is the viscous resistance of the fluid equal to $F=6\pi\eta av$ where η is the coefficient of viscosity, a the radius of the particles, and v its velocity. If X is the unbalanced force component due to the molecular impacts, then

$$m.d^2x/dt^2\!=\!-\mathrm{F}\!+\!\mathrm{X} \ m.d^2x/dt^2\!=\!-6\pi\eta adx/dt\!+\!\mathrm{X} \ m.x.d^2x/dt^2\!=\!-6\pi\eta axdx/dt\!+\!\mathrm{X} x.$$

If the mean is taken over a period of time sufficient to include a large number of impacts then $\Sigma Xx=0$ and this can therefore be neglected in performing the integration. In order to simplify substitute

$$z=d(x^2)/dt$$
 and $v=dx/dt$

giving

$$z=2xdx/dt$$
 and $dz/dt=2xd^2x/dt^2+2(dx/dt)^2$.

When these substitutions are made we get

$$(m/2)dz/dt + 3\pi\eta az = mv^2.$$

The average kinetic energy of the particle is $mv^2/2$, and according to the equipartition theorem this is the same for all molecules and therefore equals RT/2N for a particle too. Hence

$$(m/2)dz/dt + 3\pi\eta az = RT/N.$$

As

The solution of this equation for finite values of the time is, to a first approximation,

$$z = RT/3\pi\eta aN$$
.
 $d(x^2) = zdt$ then $x^2 = zt$,

hence $z=x^2/t=RT/3\pi\eta aN$.

Thus the measurement of the displacement x during the time t enables N to be calculated since the remaining quantities can be measured.

Determination of N by various methods

A number of other methods have been devised for the determination of Avogadro's number, and although most of these give only approximately correct values, one of them gives the most precise value known. These methods are as follows:—

(1) Measurements made by direct counting of radioactive particles. (This will be discussed later.)

(2) Determination of the mean wavelength of light scattered by the molecules in the atmosphere (the blue of the sky) leads to a value for N.

(3) N can be derived from the co-volume term in Van der Waal's equation of state for a gas.

(4) Studies of the opalescence of gases at their critical points can lead to a value for N (approximate).

(5) By measuring the surface tensions of monomolecular films with high precision the volumes of individual molecules can be determined and from these N can be derived with a considerable degree of accuracy.

(6) Since the charge on the electron is known N can be obtained with very high precision from electrolysis measurements, thus. The Faraday F is that quantity of electricity which liberates a gram atom of hydrogen in electrolysis. This quantity of hydrogen contains N atoms each of which carried over a charge e. Clearly F=Ne. The value of F is 96487 coulombs. This leads to

$$N = 6.023 \pm 0.001 \times 10^{23}$$
.

This value is now adopted as the most probable value for N.

The ratio of the mass of the electron to the mass of the hydrogen atom

Since Avogadro's number is the number of atoms in a gram atom, the mass of the individual hydrogen atom is obtained directly from it. It is equal to

$$M = 1.674 \times 10^{-24} \text{ gm}.$$

Since the mass of the electron is $9\cdot1066\times10^{-28}$ gm. the ratio of the two masses is

m/M = 1/1837.

REFERENCES

[&]quot;The Conduction of Electricity through Gases." J. J. Thomson and G. P. Thomson. Vol. I (1928), Vol. II (1933).

"The Electron." Millikan. (1933.)

CHAPTER 4

POSITIVE RAYS AND THE MASS SPECTROGRAPH

Positive rays

It was first noticed by Goldstein in 1886 that if the cathode of a discharge tube was perforated, rays passed through the aperture, appearing as a fine streamer of glow on the far side of the cathode. These were originally called "canal rays," but as later work showed that they are streams of positively charged particles they are now called "positive rays." The origin of these rays is quite clear. In a low pressure discharge tube the positive ions move rapidly towards the cathode, and if this is perforated some will shoot through the opening and form a pencil of ions on the far side. It was found that this ion beam has very strong ionising power, easily exciting ionisation by collision in the rarefied gas behind the cathode.

In order to throw light upon the nature of the ions constituting the positive ray beam, Sir J. J. Thomson in 1911 devised a method for measuring the ratio of the charge E to the mass M. The experimental difficulties were great, largely because of the ionising power of the rays. The method of Thomson has been developed by Aston and has resulted in fundamental discoveries of great importance concerning the ultimate structure of matter. In the following section we shall proceed to consider Thomson's method for the determination of E/M for positive rays.

Thomson's parabola method for E/M

The apparatus used by Thomson is shown in Fig. 4.1. The object of the experiment is the observation of the simultaneous deflection of the positive rays by means of parallel electric and magnetic fields. The pressure in the deflection observation chamber must be kept very low, otherwise collision effects introduce complications, but the pressure in the discharge tube

wherein the positive ions are created must be high enough to give a sufficiently intense beam of positive rays for observational purposes. The difficulty concerning these mutually contradictory conditions is overcome in the following ingenious manner.

The positive ions are created in the flask A which has a capacity of 2 litres, since a large volume steadies the discharge. The cathode B, faced with aluminium to reduce sputtering, is pierced, the diameter of the hole being 0·1 mm. This hole is continued as a fine tube, passing through a thick iron tube. Overheating of the cathode is prevented by the water-jacket C. Near the anode D is a capillary tube E, which allows a slow

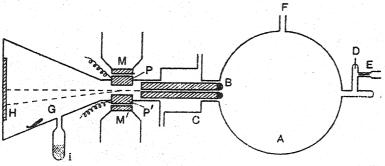


Fig. 4.1

leak of gas into the flask, the pressure in the latter being maintained fairly constant by adjusting the rate of pumping from F. A very narrow pencil of positive rays emerges from the bored cathode and then passes the pieces of soft iron PP', between which a deflecting electrostatic field X can be applied. However, PP' also constitute the pole-pieces of an electro magnet MM', although electrically isolated from each other by thin insulating sheets of mica. After passing through the electric and magnetic fields the rays then enter the very highly evacuated chamber G, called the camera, and are recorded on a photographic plate H. Owing to the narrow bore and length of the hole in the cathode, the pressure in G can be kept very low by continuous pumping with charcoal in liquid air at I, in spite of the relatively high pressure in A. By this means collision effects are avoided.

When no fields are applied, the positive rays form a central undeflected spot upon the photographic plate. The fields, being parallel, cause deflections of this spot at right angles to each other, the electrostatic deflection being in the plane of the paper. Let x be the electrostatic deflection and y the magnetic deflection, then, as already shown, if v is the velocity of the particles we get.

$$x = \frac{E}{Mv^2} \int \int_0^x (Xdx)dx$$

and

$$y = \frac{E}{Mv} \int \int_0^x (H dx) dx.$$

Since the integrals depend only upon the field strengths and the dimensions and geometry of the apparatus, we have

$$x=k_1\text{EX}/\text{M}v^2$$
 and $y=k_2\text{EH}/\text{M}v$

in which k_1 and k_2 are constants. Squaring the expression for y and dividing by that for x gives

$$y^2/x = k_2^2 H^2 E/k_1 XM$$
.

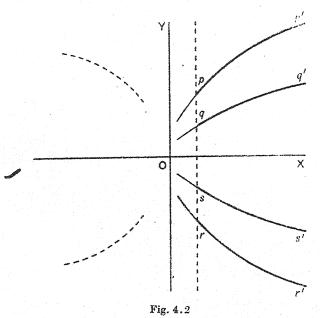
For any given values of the fields H and X the factor $k_2^2H^2/k_1X$ is constant, thus we have $y^2/x=KE/M$.

The constant K can be evaluated from measurements of the field strengths and by measuring the deflections x and y the ratio of the charge to the mass can be determined.

In the first experiments hydrogen was introduced into the tube and with this gas the value found for E/M was 9,571 e.m.u. This is the same as that found for the hydrogen ion of electrolysis, and is therefore proof that the positive rays in hydrogen consist of hydrogen atoms from which a single electron has been removed. The ions thus have each a positive charge numerically equal to that of the electron. We now know that the hydrogen atom is constituted of a relatively massive nucleus around which rotates a single electron. When this outer electron is removed we are left with a single positively charged particle (positively charged since the complete atom is neutral). The name proton has been given to this particle.

The masses of positive ions

If all the constituent ions in a positive ray beam have the same mass and charge but move with different velocities, the locus of the points formed by the rays upon the photographic plate will be a parabola, since then y^2/x is constant. The appearance on the plate will be as pp' in Fig. 4.2. If now other rays are introduced, the ions of which have the same charge as before, but with a greater mass M', a different parabola will result, and as the magnetic displacement is less



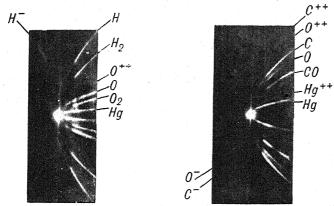
for the heavier particles the second parabola will be lower than the first. Since pp' is the parabola for hydrogen the atomic mass M' can be directly derived. The line OX cannot actually be identified upon the plate so in practice the magnetic field is reversed after half the exposure, resulting in a pair of mirror-image curves ss' and rr'. Clearly $M'/M = (pr/qs)^2$ and the latter can be measured upon the plate. This ratio therefore gives the atomic weight M' in terms of that of the hydrogen ion M.

The value of the x displacement, due to the electrical field, is inversely proportional to the kinetic energy of the constituent particles $\frac{1}{2}Mv^2$. The particles can only acquire energy up to VE where V is the potential drop across the discharge tube in which the positive ions are created, hence the parabolæ will not go through the origin but stop abruptly at a certain value of x corresponding to the maximum energy. Exceptions to this will be considered later. The type of pattern observed in general is that shown in Plate IA. It will be noticed that there are some weak curves on the left-hand side of the plate. These exist because some of the particles in their passage through the pierced cathode collide with and capture electrons. Becoming negatively charged, they suffer deflections in the opposite quadrant to those of the normal ions. They are usually found when electronegative elements such as carbon, oxygen, chlorine, etc., are present.

Frequently more than one electron is torn from an atom in the electric discharge, a multiply ionised particle being created. Such a particle has a charge which is some integral multiple of E and will therefore suffer deflections different from those for a singly ionised particle of the same mass. Consider a doubly charged ion first. If this double charge is retained the whole time, the deflections the particle will suffer will be proportional to 2E/M and are thus the same as those produced upon a particle with charge E and mass M/2. For instance, if oxygen is introduced into the discharge tube a parabola due to a mass 16, 0+, is observed and this is always accompanied by a weaker parabola corresponding to mass 8, this being the curve due to O++ ions. This type of phenomenon is fairly general and a number of examples are shown in Plate IA, in which the plus signs have only been placed against multiply ionised ions. Multiple ionisation is frequent, and in some atoms as many as eight electrons can be lost.

Amongst those multiply ionised atoms produced in the main discharge, and thus having more energy than the normal ions, are some which recapture electrons in the narrow tube. On passage through the fields they are deflected along the normal parabola if they retain a single positive charge. As they have energy in excess of VE (i.e. up to nVE where n is the degree of multiple ionisation) they extend this parabola beyond the

PLATE I



IA. Positive ray parabolas (after Aston).

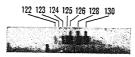


64 66 67 68 70

Zinc.

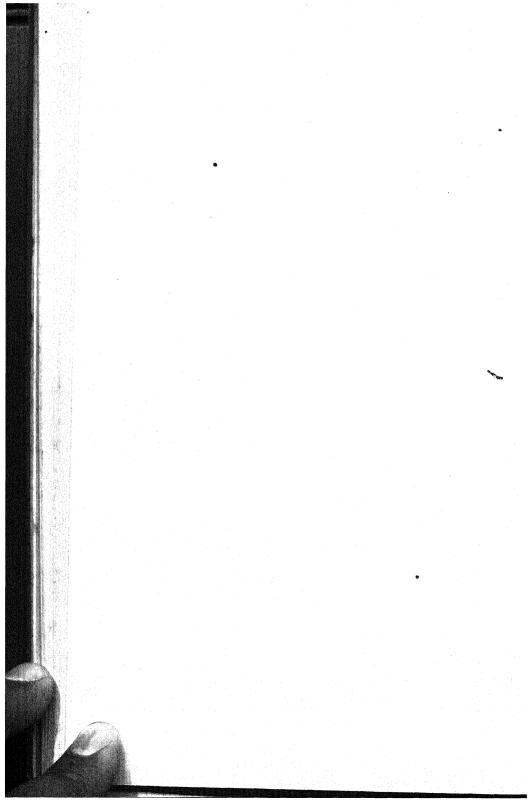


Germanium.



Tellurium.

Ic. Mass spectra by Bainbridge (after Aston).



normal x limit towards the origin. These faint extensions are visible upon the Plate IA.

The masses of the positive rays are atomic or molecular masses. By measuring these, a certain amount of new chemical information has been obtained. Thus, for instance, if a hydrocarbon, such as methane, is introduced into the apparatus, parabolas are obtained which can only be interpreted as being due to the compounds CH, CH₂, CH₃. Other chemically unstable compounds such as NH, OH, etc., can also be observed. This shows that these molecules can exist for a definite short period in the discharge tube.

Isotopes and the mass spectrograph

In 1912 Thomson examined the positive ray parabolas formed when neon was used as the source of the ions. chemical determination of the atomic weight of neon is 20.200 (based upon the scale 0=16) and, as expected, a strong line was found at a position corresponding to the atomic weight 20. This had, however, invariably associated with it a weaker parabola corresponding with atomic weight 22. No such element was known to exist. The intensity ratio of the two curves was absolutely invariant, no matter what were the conditions of discharge and pressure. The observations suggested that neon could exist in two forms chemically indistinguishable but with different atomic weights. Soddy had already arrived at a similar conclusion earlier about certain radioactive elements. These he called "isotopes" since they occupy the same place in the Periodic Table. The isotopes cannot be separated by ordinary chemical means, so that the observed chemical atomic weight is the weighted mean of those isotopes present. Thus an atomic weight 20.2 for neon would arise if there were present nine times as many atoms with atomic weight 20 as those with atomic weight 22. The atomic weight of this mixture would be $\{(9\times20)+22\}/10=20\cdot2$. It therefore became imperative to refine the apparatus in order to decide whether the atomic weight of the main parabola ion was really 20.0 and not 20.2. If this could be shown it would prove that the weaker parabola was truly due to an isotope. Aston settled this important question by designing an apparatus, now called

the mass spectrograph, with which an accuracy of one part in a thousand could be reached.

The principle of the mass spectrograph

Although the parabola method has now been considerably refined and made more accurate, all such types of apparatus suffer from the fact that the spreading out of the beam into the parabola reduces the intensity so much that there is a limit to the fineness of the slit that can be used in the cathode for the production of the positive ray beam. In the mass spectrograph the positive ions with same E/M and different velocities are brought to a focal point, the great increase in intensity permitting the use of extremely fine slits. The principle of Aston's mass spectrograph is illustrated in Fig. 4.3.

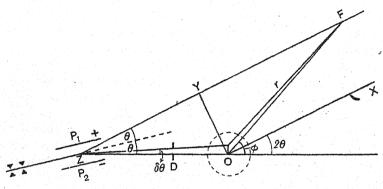


Fig. 4.3

After emerging from a pierced cathode the positive rays are made into a fine ribbon by slits, after which they pass between the electrically charged plates P_1 and P_2 . Since particles with different velocities are present, the beam is not only deflected, it is also spread out. To a first approximation the particles can be considered to radiate from Z, the mid-point of the electric field. The diaphragm D selects some of the rays allowing them to pass between the poles of an electromagnet O, the field being such that the rays are deflected in a direction opposite to that imposed on them by the electric field between the two plates P_1 and P_2 .

Let θ be the angle through which the selected rays are

deflected by the electric field and ϕ the angle through which they are deflected by the magnetic field. Then, if l is the path in the electric field X and L the path in the magnetic field H, it follows that

$$\theta v^2 = lXE/M = k_1E/M$$

 $\theta v = LHE/M = k_2E/M$.

Differentiating each of these equations gives

and $\delta\theta/\theta + 2\delta v/v = 0$ $\delta\phi/\phi + \delta v/v = 0$ thus $\delta\theta/\theta = 2\delta\phi/\phi$.

and

The small changes in angle $\delta\theta$ and $\delta\phi$ refer to the particles with the same mass and charge, but with velocities differing by δv .

Suppose all the angles are small and assume also that in effect the electric and magnetic fields act as if concentrated at points. Let ZO=b, then the width of the positive ray pencil at O is $b \cdot \delta\theta$ and at a further distance r it will be

$$W = b \cdot \delta\theta + r(\delta\theta - \delta\phi)$$
$$= \delta\theta \{b + r(1 - \phi/2\theta)\}$$

since ϕ and θ are in opposite senses.

If $\phi > 2\theta$ the width of the beam W becomes zero when $r(\phi - 2\theta) = b \cdot 2\theta$. Zero width means that the rays have come to a focus.

All the rays with common E/M focus at one point. In order to find a geometrical construction for focal points for different atoms, draw rectangular axes OX and OY so that the angle between OX and ZO is 2θ . The co-ordinates of a focal point are then $r\cos(\phi-2\theta)$ and $r\sin(\phi-2\theta)$. The latter, the Y axis co-ordinate, is, to a first approximation, equal to $r(\phi-2\theta)$ since the angle is small, and this has already been shown to be equal to $2b\theta$, i.e. it is a constant. The locus of foci is therefore a straight line parallel to OX, all the focal points for different E/M lying on this line, along which the photographic plate is placed. In effect the instrument produces mass dispersion analogous to that of an optical spectrometer and is known as a mass spectrograph.

Aston's mass spectrograph

The first instrument, built in 1919, had an accuracy of one part in a thousand, but with the later second instrument now to be described an accuracy of one part in ten thousand is attainable. As in the parabola method the positive rays are produced in a large bulb and after passing through the pierced cathode (see Fig. 4.4) traverse the slits S_1 and S_2 . These are only 0.02 mm. wide and are made by a special technique. They are 20 cms. apart and are set in line optically. The emergent pencil of positive rays is therefore very fine. The electric deflecting field is applied between the plates J_1 and J_2 which are slightly curved in order to enable them to be brought close together and yet permit the deflected beam to traverse a

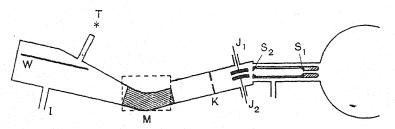


Fig. 4.4

curved path. The distance between these plates can be made as small as $1 \cdot 25$ mm.; thus large enough deflecting fields can be produced with a bank of accumulators delivering only 400 volts. The use of accumulators is of value since by this means the field can be maintained constant quite easily. The diaphragm K limits the width of the beam which is permitted to arrive at the magnetic field of the magnet M whose polepieces are curved in the shape illustrated for efficiency.

The rays are brought to a focus on the photographic plate W. The lamp T permits a fiduciary reference mark to be made upon the photographic plate. As in the parabola method it is essential to maintain as high a vacuum as possible in the camera chamber and this is done by pumping at I. Although the position of the focal line can be calculated it is found in practice that the predicted position is only approximately correct. The final focusing must be done by trial and error

and is very laborious. Special care must be taken in order to avoid sources of error which can arise if there are electrical polarisation effects upon the deflecting plates.

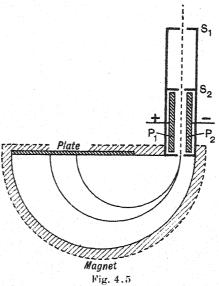
The mass ratios of ions in the discharge are obtained by comparison with a standard mass referred to a mark photographed upon the plate. For this purpose a number of empirical methods are available. In the first instances gases were employed, the behaviour of which was already known from the parabola method, and by this means lines such as C⁺⁺, O⁺⁺, C⁺, O⁺, CO⁺ were identified. These lines give effective masses of 6, 8, 12, 16, 28, etc., enabling a calibration scale to be drawn. Intermediate points can be filled in simply by moving the whole group by an alteration of the fields. If such calibration curves are to be used it is essential that the fields must be maintained constant during a photographic exposure.

The more accurate method now used by Aston is called "bracketing." The deflection equations show that if an ion of mass M appears at a particular point when the applied electric field is X, then an unknown mass M' can be brought to the same point by applying a field X' so that M'/M = X/X'. Let us now consider the ions H₂⁺ and He⁺. We shall make the provisional assumption that the ratio of the masses of these two ions is exactly 1:2. If the electric field for the hydrogen molecule is made double that for the helium atom the lines corresponding to the two ions will coincide. This is not desirable, hence three exposures are made, one for helium using a potential V upon the plates J₁ and J₂, and two for hydrogen using potentials 2V + x and 2V - x where x is a small quantity. If the ratio of the masses is exactly 1:2 the He+ line will lie exactly half-way between the two H₂⁺ lines. Any deviation from this ratio shows up as an asymmetry which can be measured in mass terms sufficiently well from a calibration curve. By bracketing unknown masses with accurately known masses great precision is attainable.

Some examples of the beautiful mass spectra which have been obtained by Aston are shown in Plate IB. The mass numbers of the isotopes are marked.

Bainbridge's method for positive ray analysis

The discoveries made by Aston with his mass spectrograph are of fundamental importance. A number of quite different types of mass spectrograph have been designed and operated by other investigators and these have all confirmed Aston's conclusions. We shall consider here one of these, namely, that designed by Bainbridge in 1930, and since improved upon. It is based upon the fact that a magnetic field alone will sort out the constituents of a mixed beam of ions according to E/M values, if all the ions in the beam have the same velocity.



Thus, if it is possible to select out all the ions from a positive ray stream with a single velocity and then to pass them through a uniform magnetic field, all those with identical E/M will traverse the same circular path and thus come to a focus. Particles with different values of E/M traverse different circular paths forming focal points corresponding to different ionic masses. The principle of the method used is illustrated schematically in Fig. 4.5.

A narrow pencil of positive ions comes down through the slits S_1 and S_2 and passes the "velocity selector" P_1P_2 . This consists of a pair of plates between which is an electric field and a magnetic field, so arranged that the deflections are in

opposite senses. Only those ions get through for which the electric and magnetic forces are equal and opposite. For these we have, if v is the velocity, X the electric and H' the magnetic field, v=X/H' (exactly as in J. J. Thomson's method for e/m for the electron).

It is clear then that all the particles which succeed in passing the selector have the same velocity. After leaving the selector the ions enter a large uniform magnetic field, the lines of force being perpendicular to the plane of the paper, and traverse circular paths, the radii of curvature being proportional to Mv/EH where H is the strength of the deflecting field. Thus the radii, and therefore the positions upon the photographic plate, are proportional to the ionic masses. The ions of different mass therefore fall upon a linear scale.

The sensitivity depends upon the area and strength of the magnetic field used to deviate the ions. Bainbridge used a very large magnet capable of giving a uniform field of 15,000 gauss over a semicircle of 40 cms. diameter. An example of the resolving power and definition attainable is shown in Plate Ic which shows the isotopes of germanium, zinc, and tellurium.

The isotopic constitution of the elements

The first success with Aston's mass spectrograph was the demonstration of the fact that there are definitely at least two isotopes in neon. Many gases were investigated and a number of metals. The production of metallic ions requires a special type of source, the best way to produce them being to heat metallic salts or crystal mixtures. The isotopic constitutions of more than seventy of the ninety-two elements have now been determined, thanks largely to the labours of Aston.

The mass standard adopted for the atomic weights of the individual isotopes is 16 for the neutral oxygen atom. The first important discovery made by Aston was that the atomic weights of all isotopes are whole numbers on this scale. The "whole number rule" is clearly of fundamental importance. The atomic weight of an element no longer has the significance previously attached to it, since in a sense it is accidental and depends upon the mixture of isotopes belonging to an element and upon their relative abundances. Chlorine is a striking

example of this, the chemical atomic weight being non-integral, 35·457. The mass spectrograph shows that chlorine has two isotopes with masses 35 and 37 the relative abundances of these two, being 76:24. The mean atomic weight of this mixture calculated from the relative abundance is exactly the chemical atomic weight.

Great complexity and variety is found amongst distributions of isotopes. For instance, some elements, like arsenic and iodine, have only a single isotope, whilst others have many, mercury possessing nine and tin eleven. Not only is there wide variation in numbers, there is a similar variation in abundances. For example, bromine has two isotopes only, the abundances of which are almost identical, whilst in oxygen there are three isotopes with relative abundance ratios about 3,000:5:1. When an element has numerous isotopes in its constitution it is always found that these occur in a fairly regular manner. For example, the isotopes of lead are 204, 206, 207, 208, or those of germanium 70, 72, 73, 74, 76. Fairly regular sequences are formed in which the increase in atomic weight in going up the series rarely exceeds two mass units for adjacent members.

Indium and tin illustrate a distribution of isotopes sometimes met with. Both of these elements have isotopes with atomic weight 115. These are therefore atoms with the same weight but occupying different places in the Periodic Table. They are called "isobars" (equal weight). Isobaric pairs occur fairly frequently and in some cases there are isobaric triplets, an example being those isotopes of tellurium, xenon, and tin which all have atomic weight 124.

The significance of the "whole number rule"

As long ago as 1815 Prout had suggested that all the elements were built up out of integral multiples of hydrogen but this was not believed because of atomic weights like that of chlorine. The difficulty about non-integral atomic weights was removed by the whole number rule. Experiment proves that every type of atom is capable of ejecting electrons under suitable conditions. Since atoms in their normal states are all uncharged they must therefore all contain units of positive electricity equal to the possible number of electrons that can

be removed. Positive ray analysis proves that the lightest observable known positively charged particle (apart from the positron and meson which are special cases that will be dealt with later) is the positive ion of hydrogen. This is the proton. Compared with the standard atomic weight taken as 16 for oxygen, the proton mass is slightly greater than unity, namely, 1.00778.

It will be shown later that experiments on the scattering of radioactive particles prove that atoms can be considered to consist of a very small massive positively charged nucleus surrounded at some distance by rapidly rotating electrons. Practically the whole mass of the atom is concentrated in the nucleus which contains the whole of the positive charge. The number of outer electrons, in the extra-nuclear or outer system, is equal to the atomic number, that is, to the numerical position the atom occupies in the Periodic Table. Hydrogen, with atomic number 1 has one outer electron, helium two, lithium three, etc. Evidence will be given later which shows that nuclei consist of protons and particles with almost the same mass but with no charge. These are the neutrons, the mass of a neutron on the same scale being 1.00895.

It is clear from the atomic weights and the nuclear charges that all nuclei (excluding that of hydrogen) must contain both protons and neutrons. For the moment let us neglect the small deviations of the masses of these particles from unity. Consider the isotopes of lithium. The atomic number of both isotopes is 3, since lithium is third in the Periodic Table. nuclear charge is therefore 3. The two isotopes have atomic weights 6 and 7. The mass and charges can only be reasonably explained by assuming that the nucleus of the lighter isotope contains three protons and three neutrons, giving a charge of 3 and a mass of 6. In like manner the heavier isotope contains three protons and four neutrons giving a charge of 3 and a mass of 7. If a proton is added to a given nucleus the charge is increased so that the atomic number rises and there results an atom in the next place in the Periodic Table. On the other hand, if neutrons are added to a given nucleus the atomic number is not affected but the mass increases and thus we have a new isotope.

The mass defect and packing fraction

With increased precision in measurement Aston found that in practically every case slight deviations from the whole number rule existed. The deviation was relatively large only in the case of hydrogen, hence 16 was still retained as the standard, rather than H=1, but this was quite arbitrary. On this basis some of the exact weights, selected at random, are as follows:

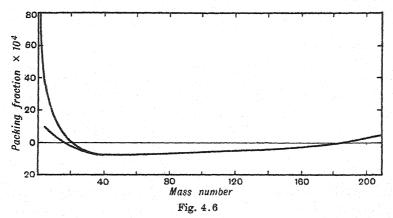
H He Li O As Ba Tl 1-00778 4-083 7-012 16 74-934 137-916 203-036

The importance to be attached to these deviations becomes apparent when the relative masses of hydrogen and helium are considered. The nuclear charge of helium is 2 and the mass about 4, and it is logical to assume that the helium nucleus is built up out of two protons (hydrogen nuclei) and The mass of this combination should be * two neutrons. 2(1.00778+1.00895), i.e. 4.03346. As the above table shows. the atomic weight of helium is considerably less, being in fact lighter by about 0.030 mass units. The difference can be explained by the Theory of Relativity, according to which mass and energy are interchangeable and can be converted into one another. (This has been experimentally confirmed.) Inside the helium nucleus the constituent particles are very tightly packed together, as proved by measurements of the nuclear radius. It can be shown that in the original process of the formation of such a nucleus energy must be released in very large amounts before a stable packing state is reached. This enormous loss of energy, which is related to the binding force between the particles, means a corresponding loss of mass. Clearly the stability of the final nucleus formed depends upon the amount of mass which has been lost in the form of radiation energy. The conditions necessary for the building up of nuclei out of their constituent particles, with the liberation of massenergy as radiation, probably exist in the centres of hot stars where temperatures of many millions of degrees are reached.

The deviation of the mass of a nucleus from the whole number is called the mass defect. It can be expressed in a

^{*} The way in which the neutron mass is measured will be described later.

convenient form thus. If N is the mass number of an isotope, i.e. the nearest whole number, the mass in general is $N-\delta$ where δ represents the defect. The quantity δ/N is called by Aston the packing fraction, since it is a measure of the amount of packing of the constituent particles. The packing fraction, which is generally expressed in parts per 10^4 , is a very important quantity, being a measure of the stability of the nucleus. When the observed packing fractions are plotted against the mass numbers of atoms they are found to fall upon the smooth curve shown in Fig. 4.6. Those light atoms with atomic masses which are exact multiples of four lie upon a separate lower spur. These nuclei are more stable than the others.



The packing fraction of the helium nucleus is such that there are strong grounds for believing that this is a particularly stable combination of nuclear particles. It is probable that the helium nucleus can act as a sub-unit in nuclear structure. For this reason many light atoms whose mass numbers are divisible by four, are very stable and lie on a separate lower stability curve.

Oxygen has been chosen as the standard for convenience. If hydrogen were taken as standard the numerical values of the packing fractions and mass defects would differ. For instance, the mass defect of oxygen would then be 0·12.

According to the Relativity Theory the equivalence of mass and energy is given by $E=mc^2$. That is, a mass m grams is equivalent to E ergs, c being the velocity of light. If a gram of hydrogen is transmuted into helium this formula shows that

 7.0×10^{18} ergs of energy will be radiated. This is approximately equal to 200,000 kilowatt hours. Clearly this enormous amount of energy is a possible source for stellar energy of radiation; this will be discussed later.

The discovery of heavy hydrogen (deuterium)

The discovery of an isotope of hydrogen with mass 2 and its production in quantity has become of great importance. particularly for chemistry. The discovery goes back to the detection of isotopes of oxygen by spectroscopic means. Aston had not been able to observe any oxygen isotopes other than 16 which was one reason for its adoption for the standard of atomic weight. The positive ray method is only one of the experimental procedures available for the study of isotopes. The differences in the masses of the isotopes constituting an element can have considerable influence upon the finer details of the spectra of molecules. Isotopes can therefore be studied by spectroscopic means. For instance, the spectrum given out by the molecules of chlorine has in it lines which can be attributed to the molecules 35Cl-35Cl, 35Cl-37Cl, and 37Cl-37Cl. The careful study of the molecular spectrum of oxygen revealed the existence of an isotope 180 present to an extent of 1 in 630 parts of ¹⁶O and a fainter ¹⁷O present to the extent of 1 in 3,150 parts of ¹⁶O. Clearly Aston's isotopic weights are all referred to the main isotope of mass 16 but the chemical atomic weights are based upon the atomic weight of the oxygen mixture. since this mixture as a whole takes part in chemical reactions. The weight of this isotopic mixture is 16.0035 on Aston's scale, hence Aston's weights have to be reduced by 2.2 parts in 10,000 to bring them to the chemical scale.

The accurately measured chemical atomic weight of hydrogen is 1.00777 and the value previously found by Aston before the discovery of the oxygen isotopes was 1.00778, based upon a value of 16 for oxygen. This agreement, although actually accidental, was considered eminently satisfactory. However, the discovery of ¹⁷O and ¹⁸O requires a reduction in the mass spectrum scale in order to bring it into line with the chemical scale and when this is done the mass spectrographic value for hydrogen becomes 1.00756. The agreement has been turned into a discrepancy of 0.00022 mass units. This could be

accounted for if it were assumed that hydrogen contained a small quantity of a heavy isotope, for then the chemical atomic weight would be that of the mixture and therefore greater than the mass spectrographic value. It requires only 1 part in 6,500 of an isotope of mass 2 to produce the alteration in weight observed by Aston.

Urey and his collaborators thereupon undertook the fractionation of liquid hydrogen, and owing to the fact that there is 100 per cent. difference in the masses of the two isotopes this method, inefficient in previous cases, was highly effective. The suspected isotope was successfully separated, being detected spectroscopically. It was named deuterium. It was later discovered that the water left in old oft-used electrolytic cells has relatively a very high concentration of deuterium (D) in the form of heavy water D₂O. This arises because of the different mobilities of the electrolytic ions of hydrogen and deuterium respectively. By electrolytic methods it is now possible to obtain practically pure heavy hydrogen or heavy water in large quantities. Litres of deuterium can now be bought commercially at a low price.

The importance of deuterium to chemistry as an indicator is very great. It is easy to replace hydrogen by deuterium in, say, a complex organic compound containing many hydrogen atoms, since chemically the two are very nearly indistinguishable. Thus a particular hydrogen bond can be marked and followed in various reactions, for the deuterium can always be recognised by its density when it is liberated or transposed. Deuterium has also proved of great value in studying the structure of nuclei both as an object for disintegration and as a projectile for the production of disintegrations.

The separation of isotopes

Since isotopes have virtually identical electronic structures their chemical properties are, for almost all practical purposes, indistinguishable (apart from the case of hydrogen). They can therefore mostly only be separated by physical methods which depend upon the atomic mass. In a broad sense these can be divided off into methods available only for the very light elements and those available for light and heavier atoms. "Chemical" methods which involve either (1) electrolysis,

(2) distillation, or (3) chemical change, can only be used with atomic weights up to that of oxygen. We shall first deal briefly with these processes.

Electrolysis

If a dilute electrolyte is electrolysed, with emission of hydrogen, the differential ionic mobility of the two isotopes of hydrogen leads to a more rapid evolution of light hydrogen. The heavy hydrogen therefore concentrates in the electrolyte in the form of heavy water. The solution is electrolysed until it is very concentrated. After dilution, the process is repeated until the original volume is reduced to one nine-thousandth part. The residue is then 90 per cent. heavy water. When the reduction is carried to 1 part in 25,000 the concentration reaches 99-9 per cent. and is thus virtually a complete isotopic separation. Some 30,000 ampere hours are required for the production of 1 gram of heavy water. This method is only effective for hydrogen.

Distillation

If the heat capacity of a liquid at low temperatures differs appreciably from that at normal temperatures, then it can be shown that there will be differences in the vapour pressures of any isotopes present. The differences in specific heat sufficing to make this property applicable to isotope separation are only exhibited by the liquefied gases hydrogen, helium, and neon. In 1935 Keesom, Van Dyk, and Haantjes succeeded in separating the neon isotopes (partially) by this method.

Exchange chemical reactions

In exchange reactions between isotopically different molecules, the equilibrium constants differ slightly from unity. For example, the reaction \$^{13}CO + ^{12}CO_2 = ^{12}CO + ^{13}CO_2\$ has an equilibrium constant of \$1.098\$ at \$273°\$ A. and \$1.029\$ at \$600°\$ A. By utilising this (and other analogous reactions) increased concentrates of isotopes, but not complete separations, can be achieved.

The above three methods are all slow and laborious and are not so strikingly successful as the methods to be described below, which can be applied to both light and heavier atoms.

In all separations the lower the initial abundance ratio, the easier it is to concentrate the less abundant. The more general methods now to be described involve (1) diffusion, (2) thermal diffusion, (3) the mass spectrograph.

Diffusion

Since the diffusion rate of atoms is inversely proportional to the square root of the mass there will be slight differences in diffusion velocities of isotopes. In 1932 Hertz successfully applied this principle to the separation of the isotopes of neon. If a gas is diffused through a porous tube a slight partial isotopic enrichment takes place. Forty-eight such diffusion tubes built into twenty-four pumps are arranged in cascade, forming a closed circuit through which the continuously enriching gases are constantly circulated by an elaborate arrangement of pumps and valves. With later improvements of the apparatus great speed in separation was achieved. For example, in 8 hours ²⁰Ne, containing only 1 per cent. of the heavier isotope, was extracted, the quantity being 55 c.c. at normal temperature and pressure.

Hertz has modified the apparatus so that the diffusion takes place through a wall of streaming vapour (e.g. mercury). Neon, argon, and even bromine isotopes have been rapidly separated by this process. The rate of transport of the material is about 1 c.c. per day with this arrangement.

Thermal diffusion

A very effective method for isotopic separation was introduced in 1938 by Clusius and Dickel who employed thermal diffusion, applied in the following way. If a mixture of gases of different molecular weights (whether differing isotopically or chemically) is placed between vertical parallel walls, one of which is hot and the other cold, the components separate out because: (1) thermal diffusion concentrates the heavier component near to the cold wall and the lighter near to the hot wall; (2) thermal convection cause the colder (and therefore heavy-enriched) gas to flow down in the neighbourhood of the cold wall whilst the hotter (and light-enriched) gas flows up near the hot wall. This leads to two gas streams which result in separation, until back diffusion leads to a stabilised distri-

bution. In the first experiments a vertical tube with a hot co-axial wire was used as the separating system. With a length of 3 metres and a temperature difference of 600° C., a mixture of gases such as bromine and helium was rapidly completely separated into its components. A 50 per cent. enrichment of the chlorine isotopes was attained. Later with a tube 36 metres long a complete separation of the chlorine isotopes (to within 0.4 per cent.) was achieved rapidly, 8 c.c. of gas being produced in 24 hours. It is found that a tube but 1 metre in length is equivalent in speed to 12 Hertz diffusion pumps in cascade.

The method is not only applicable to gases but can also be used with liquids. With zinc chloride the zinc isotopes 64 and 68 can be partially separated. The thermal diffusion process is by far the simplest and most effective method yet devised. It is, however, limited to materials which are not affected by the high temperatures used and which will not react with the hot wire. Moreover, the process is more complex than has been indicated and its failure to separate components effectively in certain cases has to be attributed to peculiarities in the molecular force-fields for the molecules concerned.

The mass spectrograph

It is clear that the traces made by isotopes when the mass spectrograph is used represent a complete separation of isotopes. If the ion sources used are very intense, the different isotopes can be collected separately and removed for examination. At present the quantities which have been obtained by this method are very minute, but as the separation is complete the process is useful for special cases.

REFERENCES

[&]quot;Mass Spectra and Isotopes." Aston. (1942).
"Reports on Progress in Physics." Vol. 4 (1940). Physical Soc. of Lond.

CHAPTER 5

THE QUANTUM THEORY

Difficulties with the classical theory of radiation

Towards the end of the nineteenth century an anomalous situation had arisen in connection with the theory of "black body" radiation. By definition a "black body" is one that completely absorbs all the radiation falling upon it, or conversely behaves as a perfect radiator when heated. The classical electromagnetic theory of heat radiation can be applied to the calculation of the density of the radiation in an enclosure surrounded by a black body maintained at a uniform temperature. Such a radiator will, on classical theory, emit energy radiation of continuously variable wavelength. Rayleigh and Jeans have derived an expression giving the radiation energy per unit volume of the ether in equilibrium in such an enclosure, in a wavelength range extending between λ and $\lambda-d\lambda$.

In Jeans' derivation of this radiation law the radiation is imagined as broken up into monochromatic wave trains and the number of such trains, or equivalent degrees of freedom, lying between the wavelengths λ and $\lambda - d\lambda$ is determined. The energy carried by each degree of freedom is known from general statistical theory, hence the energy density of the radiation can be obtained.

Suppose radiation of wavelength λ forms steady stationary waves by being enclosed between two perfectly reflecting walls separated by a distance x. Then $x=n\lambda/2$ where n is an integer. Consider now a cubical box with perfectly reflecting walls of side a enclosing radiation which produces stationary waves. Then

$$la = n_1 \lambda/2$$

$$ma = n_2 \lambda/2$$

$$na = n_3 \lambda/2$$

where n_1 , n_2 , n_3 are integers and l, m, n are the direction

cosines for the radiation direction—since only when these conditions are satisfied will the corners of the cube remain permanently nodes of vibration. Square and add, giving

$$a^2(l^2+m^2+n^2)=\lambda^2(n_1^2+n_2^2+n_3^2)/4$$
 since
$$l^2+m^2+n^2=1$$
 then
$$(n_1^2+n_2^2+n_3^2)/a^2=4/\lambda^2=4\nu^2/c^2$$
 hence
$$\nu=(n_1^2+n_2^2+n_3^2)^{\frac{1}{2}}c/2a$$

where ν is the frequency and c the velocity of the radiation.

The number of possible stationary vibrations which occur between the frequency limits ν and $\nu+d\nu$ corresponding to the

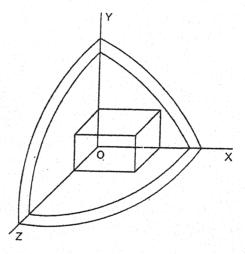


Fig. 5.1

wavelengths λ and $\lambda-d\lambda$, can now be found by means of a graphical construction. Using rectangular axes (Fig. 5.1) the values $n_1c/2a$, $n_2c/2a$, $n_3c/2a$ are plotted as the x, y, z coordinates. This gives a cubic lattice of points in which the distance of a point from the origin is (from above) equal to ν . The total number of combinations of n values such that

$$\nu < (n_1^2 + n_2^2 + n_3^2)^{\frac{1}{2}}c/2a < \nu + d\nu$$

is the number of points lying in one octant between the spheres of radii ν and $\nu + d\nu$. The volume of the complete spherical shell is $4\pi \nu^2 d\nu$. The volume of the elementary cubic lattice corresponding to one vibration is that given by $n_1 = n_2 = n_3 = 1$,

i.e. it is $(c/2a)^3$. Hence the total number of independent vibrations in the cube of volume a^3 is

$$4\pi \nu^2 d\nu / 8(c/2a)^3$$
.

The number of vibrations per unit of volume is then

$$4\pi v^2 dv/c^3$$

—taking no account of polarisation. However, we should consider standing waves of mutually perpendicular polarisation as distinct modes of vibration. The above number must therefore be doubled, hence the total number of vibrations, or equivalent wave trains is

$$8\pi\nu^2 d\nu/c^3$$
.

An ether vibration, or wave train, may be considered to be a dynamical system in which the average kinetic energy is equal to the average potential energy, both being equal to kT/2 where k is Boltzmann's constant (Avogadro's number into R, the gas constant) and T the absolute temperature. The total energy for each mode of vibration being kT the radiation density is therefore

$$d\mathbf{E} = 8\pi v^2 dv k \mathbf{T}/c^3 = 8\pi k \mathbf{T} \lambda^{-4} d\lambda \qquad . \qquad . \qquad . \qquad (1)$$

This is the Rayleigh-Jeans classical radiation formula.

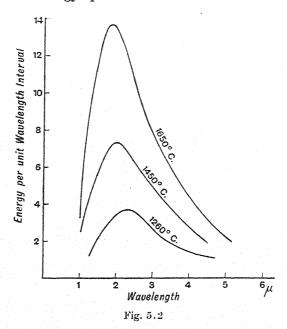
According to this expression, the energy radiated (by a black body) in a given wavelength range $d\lambda$ increases rapidly as λ becomes smaller and in fact approaches infinity with small enough wavelength. The experimentally observed radiation curve is in complete disagreement with this conclusion. Furthermore, the energy carried by all wavelengths (continuous emission) is

$$\int\!\! d\mathbf{E} \!=\! \! \int_0^\infty \! 8\pi k \mathrm{T} \lambda^{-4} d\lambda$$

where λ goes from zero to infinity. The integral is equal to infinity for any value of T other than T zero, which interpreted physically means that the total energy radiated per unit of time per unit area is infinite at all temperatures. Clearly this is an entirely false conclusion. Both conclusions are incorrect and both fail completely to account for the observed dependence of radiation on temperature. The Stefan-Boltzmann law, which has been confirmed in the laboratory, shows that heat radiation is proportional to T^4 and the observations of Lummer

and Pringsheim on black body radiation show that the Rayleigh-Jeans law is not obeyed.

A black body does not exist in nature but an equivalent can be devised, since it is axiomatic that the radiation coming out of the small opening of an almost completely elosed uniformly heated hollow enclosure is equivalent to black body radiation. In 1897 Lummer and Pringsheim, by means of a radiation bolometer, measured the wavelength distribution in the emitted energy spectrum of such a black body radiator.



Radiation curves were measured for different emission temperatures between 1,000° C. and 1,650° C. The curves found are shown in Fig. 5.2. Actually the Rayleigh-Jeans formula can be made to fit the curves for sufficiently large values of λ , but for small λ extreme disagreement is obvious. The amount of energy radiated at any temperature is proportional to T⁴, which corresponds to Stefan's law, and is not, of course, infinite. Since the mathematical derivation of the classical formula is free from error, the contradiction with experiment can only mean that the fundamental assumptions used in deriving the expressions are at fault.

Difficulties with the classical theory of specific heats

According to the kinetic theory, which is based upon statistical and thermodynamic considerations, the mean energy of each degree of freedom of a gas molecule at temperature T is kT/2, k being Boltzmann's constant. For a monatomic gas there are three degrees of freedom for translational motion so that the total average energy per particle is 3.kT/2. A solid is supposed to consist of individual atoms which are partially bound to their neighbours, but yet can resonate and vibrate fairly freely when heat is supplied. There will be N atoms in a gram molecule so that the kinetic energy of oscillation for this amount will be 3kTN/2, which equals 3RT/2 since k=R/N. The mean potential energy for each of the three components of vibration for each particle will also be kT/2 (it can be proved that the average potential energy equals the average kinetic energy in S.H.M.). Hence the total energy is E=3RT. By definition the specific heat at constant volume is $(\delta E/\delta T)_v = 3R$. From this reasoning the specific heats of all solids should be the same and independent of temperature.

The numerical value of 3R is 5.96 calories per gram mol and the above treatment was considered to be a theoretical vindication of the early law of Dulong and Petit, which states that the heat capacity of the gram mol of most solid elements is approximately 6. However, this law fails completely when applied to light elements like carbon, and breaks down generally at low temperatures, as all specific heats tend to zero at the absolute zero of temperature. This failure cannot be accounted for by the classical theory of heat absorption which requires a body to be able to absorb heat continuously in indefinitely small amounts, so that again it is apparent that the fundamental assumptions of the classical view are at fault.

The quantum theory

In the failure of the classical methods to account for the radiation curves and for the specific heat data, the fundamental assumptions of the classical view must be to blame, and in 1900 Planck proposed a revolutionary hypothesis by the aid of

therefore

which he deduced the true law of heat radiation. This, known as the quantum theory, has profoundly altered the whole outlook of physics, which up to the end of the nineteenth century had been developed on the basis of continuity in energy. Planck's radical change amounted to the introduction of the concept of atomicity in the realm of energy. He made four basic assumptions, namely:

(1) A black body contains simple harmonic oscillators which can vibrate with all possible frequencies.

(2) The frequency radiated by an oscillator is the same as the frequency of kinetic motion.

(3) The emission of radiation takes place at separate intervals, the amplitude remaining constant in the intervening periods.

(4) An oscillator emitting a frequency ν can only radiate in units, or quanta, of the magnitude $h\nu$, where h is a universal constant.

The fourth assumption is most revolutionary in nature, assuming as it does that energy can only be radiated in discrete quantities or packets, and not in continuously variable amounts. The constant h is now called Planck's constant.

The quantum theory of heat radiation

Consider a black body radiator containing a number of linear oscillators which can only vibrate with the integral energy values 0, $h\nu$, $2h\nu$, $3h\nu$, . . . $nh\nu$. . . Let there be N_0 vibrators present with zero energy. From the kinetic theory of equipartition of energy, the number of vibrators with energy ϵ is given by $N_0e^{-\epsilon/kT}$, hence the number with energy $h\nu$, $2h\nu$, $3h\nu$. . . $nh\nu$. . . respectively is N_0e^{-x} , N_0e^{-2x} , N_0e^{-3x} . . . N_0e^{-nx} . . . where $x=h\nu/kT$. Summing gives N, the total number of resonators of the given fundamental frequency present, so that

$$N = N_0 + N_0 e^{-x} + N_0 e^{-2x} + N_0 e^{-3x} \dots N_0 e^{-nx} \dots (2)$$

= $N_0 (1 + e^{-x} + e^{-2x} + e^{-3x} \dots e^{-nx} \dots)$

the series is equal to the expansion of $(1-e^{-x})^{-1}$ so that

The total amount of energy associated with the N_0e^{-x} resonators each with energy $h\nu$ is $h\nu$. N_0e^{-x} , and similarly that

associated with the resonators of energy $2h\nu$ is $2h\nu$. N_0e^{-2x} , etc. Summing the energy for all the resonators of frequency ν present gives

$$E = O.N_0 + h\nu.N_0e^{-x} + 2h\nu.N_0e^{-2x} + 3h\nu.N_0e^{-3x}...nh\nu.N_0e^{-nx}$$

$$= h\nu.N_0e^{-x}(1 + 2e^{-x} + 3e^{-2x}...ne^{-(n-1)x}...)$$

$$= h\nu.N_0e^{-x}(1 - e^{-x})^{-2}$$

$$= h\nu.N_0e^{-x}/(1 - e^{-x})^2$$

substituting $N=N_0/(1-e^{-x})$ from equation (3) gives

$$E = h\nu \cdot Ne^{-x}/(1-e^{-x})$$

= $h\nu \cdot N/(e^x-1)$.

Finally, replacing x by $h\nu/kT$ we get

$$E = h \nu N/(e^{h\nu/kT} - 1)$$
 (4)

From this expression the average energy per resonator of frequency ν is $E'=h\nu/(e^{h\nu/kT}-1)$ (5)

To obtain Planck's energy distribution law the energy per resonator must be multiplied by what is effectively the number of resonating modes of that frequency per unit volume of the ether. This will give the energy in a unit volume. The number of modes of vibration of the ether lying in the wavelength range . . . λ — λ - $d\lambda$ is equal to $8\pi\lambda^{-4}d\lambda$, hence the distribution law becomes

$$d\mathbf{E} = 8\pi \lambda^{-4} d\lambda . h \nu / (e^{h\nu/kT} - 1) \quad . \quad . \quad . \quad . \quad (6)$$

or in terms of λ

$$d\mathbf{E} = 8\pi hc\lambda^{-5}d\lambda/(e^{hc/k\lambda T}-1)$$
.

When λ is large, or more precisely when $h\nu$ is small compared with kT, the factor

$$(e^{h\nu/kT}-1)$$
 tends to $h\nu/kT$.

It will be seen from (5) that in this case the average energy E of one degree of freedom becomes equal to kT (and independent of frequency), which is the value derived classically. For this reason the distribution law derived from the quantum theory will give the same numerical result as that derived classically, when long wavelength radiation is considered.

If Planck's radiation formula is written as

$$d\mathbf{E} = (8\pi k \mathbf{T} \lambda^{-4} \cdot d\lambda) \cdot \{(h\nu/k\mathbf{T})/(e^{h\nu/k\mathbf{T}} - 1)\}$$

the first bracket term is equivalent to the Rayleigh-Jeans formula, from which it can be concluded that the quantum theory formula can be derived from the classical formula by

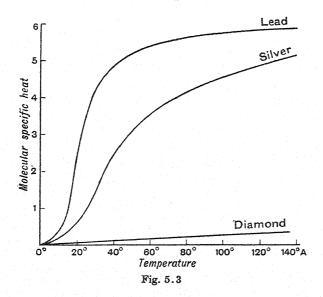
multiplying the latter by a factor $(h\nu/kT)/(e^{h\nu/kT}-1)$ which can be called Planck's multiplying factor.

Planck's law fits the experimental radiation curves very closely indeed. Differentiating shows that the energy per unit wavelength interval has a maximum value for a given wavelength (T constant), which is what is observed in the experimental curves in Fig. 5.2. Wien had earlier postulated his "Displacement Law" which stated that the product of the temperature, and the wavelength of the maximum corresponding to this temperature, is constant. From Lummer and Pringsheim's curves the wavelengths of the maxima at different temperatures can be measured and so the constant can be determined. It is thus found experimentally that $\lambda_{\text{max}}T = 0.294$. If Planck's equation is differentiated with respect to λ (maintaining T constant and assuming the currently accepted values for h, k and c) it is found that a maximum for λ occurs at $\lambda T = 0.288$, which is in excellent agreement with the observed value, when experimental errors are taken into account. It will be seen that the quantum theory of heat radiation succeeds completely where the classical theory breaks down. It may therefore be expected that quantum considerations will lead to a solution of the specific heat difficulties. This is indeed the case.

The quantum theory of specific heats

Dulong and Petit's law for the specific heat of solids fails at low temperatures and with low atomic weights. This is shown clearly by the curves of variation of specific heat with temperature (it will be remembered that the classical theory predicts that the specific heat is independent of temperature). The specific heat curves for a number of materials are shown in Fig. 5.3. The general properties of the curves are as follows. At ordinary and high temperatures the specific heats converge to 3R but tend towards zero at very low temperatures. For a short way up from the absolute zero the specific heats are accurately proportional to T3. This is called Debye's law and is generally applicable. Of great importance is the fact that all the curves for different materials have exactly the same shape and can be made to coincide by altering the scale of the temperature axis. This means that $C_v = f(\theta/T)$ where f is the same function for all substances and θ is a temperature characteristic of each substance. Thus if θ_1 and θ_2 are the characteristic temperatures for two materials, and if T_1 and T_2 are the temperatures at which the respective values of C are the same for both, then it follows from the identity of the functions that $\theta_1/T_1 = \theta_2/T_2$.

Since the classical theory fails, Einstein invoked the quantum theory in 1907 in order to attempt an explanation. He suggested that by analogy with the radiation formula, the quantum theory expression for specific heat could be obtained



if the classical value for the energy of each particle was multiplied by Planck's factor. The classical energy value for a gram mol is 3RT so that, according to Einstein, the energy given by the quantum theory is

$$\mathbf{E} = 3\mathbf{R}\mathbf{T}(h\nu/k\mathbf{T})/(e^{h\nu/k\mathbf{T}}-1).$$

If T is large compared with $h\nu/k$ (which can be thought of as the characteristic temperature θ) this tends to 3RT, *i.e.* to the classical value. If, however, T is not large, then the expression for the specific heat will be obtained by differentiating, giving

$$C_v = (\delta E/\delta T)_v = 3R \cdot e^{h\nu/kT} (h\nu)^2 / \{kT(e^{h\nu/kT} - 1)\}^2$$

This equation for the specific heat gives temperature curves very similar to those found experimentally but it fails at low temperatures approaching the absolute zero. The disagreement is due to the neglect of the mutual forces exerted by the atoms upon each other. It was left to Debye in 1912 to develop a specific heat formula which gives excellent agreement with experiment over the whole observable temperature range.

Debye's theory of specific heats

In Debye's theory a solid is assumed to be an elastic body, the vibrations of the whole being considered. From the theory of vibrations it can be proved that the number of independent vibrations per unit volume, which lie in the frequency range between $\nu + d\nu$ and ν is $4\pi(1/c_i^3 + 2/c_i^3)\nu^2 d\nu$ where c_i is the velocity of propagation for longitudinal waves and c_i that for transverse waves. It will be noticed that three independent types of vibration can be transmitted, one longitudinal and two transverse (longitudinal waves cannot be polarised, but transverse waves can), hence the factor 2 in the term involving c_i . If V is the volume of a gram molecule of the solid the number of vibrations for this amount of material will be

$$4\pi V(1/c_l^3+2/c_l^3)v^2dv$$
 (7)

From the fundamental quantum theory, the average energy (see equation (5)) for a vibration of frequency ν is equal to $h\nu/(e^{h\nu/kT}-1)$, hence the energy in the solid for frequencies between $\nu+d\nu$ and ν is

$$dE = 4\pi V(1/c_l^3 + 2/c_l^3)h\nu^3 d\nu/(e^{h\nu/kT} - 1) \quad . \quad . \quad (8)$$

If the solid were considered to be continuous the possible values that could be given to ν would be infinite and it would be necessary to integrate from zero to infinity in order to obtain the total energy associated with the body. But here Debye takes into account the atomic structure of matter by assuming that the frequencies can only reach a definite upper limit ν_m . This limit is chosen so that the total number of possible independent vibrations is equal to the number of vibrations of the separate atoms in the solid. If there are N_0 atoms in the volume V, the number of vibrations possible will be equal to $3N_0$ because each atom has three directional degrees of freedom. However, the total number of vibrations is also the integral of equation (7), hence we have

$$3N_0 = \int_0^{\nu_m} 4\pi V(1/c_l^3 + 2/c_l^3) \nu^2 d\nu$$

which gives

$$3N_0 = \frac{4\pi}{3} V(1/c_l^3 + 2/c_l^3) \nu_m^3$$

therefore

$$(1/c_l^3 + 2/c_l^3) = 9N_0/4\pi V \nu_m^3$$
 (9)

Substituting in equation (8) gives

$$dE = 9N_0h\nu^3d\nu/\nu_m^3(e^{h\nu/kT}-1).$$

Integrating for the total energy associated with the solid,

$$E = \frac{9N_0}{\nu_m^3} \int_0^{\nu_m} h \, \nu^3 d \, \nu / (e^{h\nu/kT} - 1) \qquad . \qquad . \qquad (10)$$

If now equation (10) is differentiated with respect to T, the following expression is obtained for the specific heat

$$C_v = (\delta E/\delta T)_v = \frac{9N_0}{\nu_m^3} \int_0^{\nu_m} h^2 \nu^4 e^{h\nu/kT} d\nu / \{kT^2 (e^{h\nu/kT} - 1)^2\}.$$

This can be integrated by parts and then evaluated graphically. It can be rewritten as

$$C_v = \frac{9N_0k}{(h\nu_m/kT)^3} \int_0^{\nu_m} (h\nu/kT)^4 e^{h\nu/kT} d(h\nu/kT)/(e^{h\nu/kT} - 1)^2$$

from which it will be seen that the right-hand side of the expression is a function only of $h\nu_m/kT$, hence

$$C_v = f(\hbar \nu_m/kT)$$

= $f(\theta/T)$ where $\theta = \hbar \nu_m/k$.

The first point of interest is that the above law for the specific heat is exactly that deduced empirically from the experimental curves in Fig. 5.3. The quantity θ only involves the universal constants h and k and the frequency ν_m which depends upon the particular substance under consideration. When the Debye function is integrated numerically, say for copper, the curve which is obtained, shown in Fig. 5.4, gives an exact fit with the experimentally determined values of the specific heat. The value of θ for any substance can be calculated from specific heat measurements. However, from equation (9)

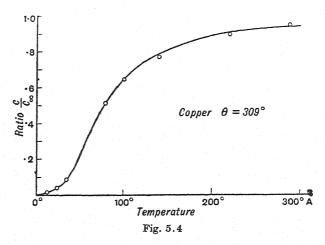
$$\nu_m^3 = 9N_0/4\pi V(1/c_l^3 + 2/c_l^3)$$

hence θ can be independently calculated from a knowledge of the elastic constants of the material. The values so found are in complete agreement with those obtained thermally.

When T is small, the Debye equation becomes approximately $C_n = \pi^4 \cdot 12R \cdot T^3/5\theta^3$

which is the theoretical origin of Debye's T³ law for low temperatures.

Strictly speaking the above Debye theory only applies to isotropic bodies in which all the atoms are identical but it can be extended so as to include crystals and chemical compounds. For non-isotropic bodies three terms have to be used, each with a different θ value, corresponding to the three principal axes in the body. In the case of crystals account is taken of



the lattice structure by introducing three θ terms and also a number of others arising from individual frequencies. In effect the three Debye terms correspond to the vibration of the molecules as a whole, but in addition terms have to be added representing the vibrations of the individual atoms in the molecule. Since the latter act as simple oscillators, these additional terms will be identical with those derived by Einstein in his earlier specific heat formula. Formulæ which include both Debye and Einstein terms have been found to fit practically all the observed specific heat data of chemical compounds and crystals.

The photo-electric effect

A detailed discussion of the photo-electric effect will not be given here, but it will be briefly considered at this point since

its explanation by Einstein was an important step in the early development of the quantum theory. When either light or X-radiation is allowed to fall upon metallic surfaces, electrons are ejected from the metals. For any particular surface irradiated by light of different frequencies it is found that the kinetic energy of the ejected electrons does not depend upon the intensity of the light but varies linearly with the frequency. This can be expressed as $mv^2/2 = hv - P$ where $mv^2/2$ is the kinetic energy of the electron, ν the frequency of the radiation, h a constant. P is the amount of work required to extract the electron from the metal and varies from material to material. The constant h has exactly the same value as Planck's radiation constant and must be identified with this. Thus when a quantum of energy $h\nu$ separates an electron from a metallic surface, work has to be done and the surplus energy

reappears as the kinetic energy of the electron.

The intensity of the irradiating beam has absolutely no effect upon the energy of the individual emitted electrons, only varying the total number which is ejected. Experiments carried out with very weak beams of light showed that the photo-electric effect commences immediately a metal is irradiated (or at least within 3×10^{-9} second). Einstein pointed out that in some of these experiments the beam used was so feeble that many hours of illumination would be required before an amount of energy equal to $h\nu$ could fall upon a single atom, yet in spite of this the effect is instantaneous. Einstein's calculation was based upon the assumption that the energy is uniformly distributed over the spherical wave-front of the beam, as the classical theory demands. Since the photo-electric emission is instantaneous, and as a quantity of energy $h\nu$ is definitely required for the ejection of an electron with kinetic energy $mv^2/2$, Einstein was forced to conclude that the energy $h\nu$ is not propagated as a wave but rather more in the nature of a particle travelling in a single direction. There is no sideways spreading so that the complete individual quantum of energy can be absorbed by any atom it may strike, no matter how far the quantum has travelled. The light quantum, or photon, as it is called, is virtually like a projectile travelling with the velocity of light.

It will thus be seen that the photo-electric effect affords

decisive evidence in support of the quantum theory and suggests an extension of Planck's original hypothesis. For not only is energy radiated in discrete amounts, but these individual amounts travel as separate packets of energy rather than in the form of spreading out wave pulses. The absorption of a quantum of radiation can now be more easily understood, since radiation energy impinges upon an absorbing surface in the form of a shower of quanta and must clearly be absorbed in integral amounts.

The hydrogen spectrum

It has long been known that hydrogen, which, being the lightest atom, is the least complex in its inner structure, emits a line spectrum which exhibits definite regularity. The visible hydrogen line spectrum consists of four lines forming a regular series which gradually closes up. This series actually extends into the ultra-violet region, the lines becoming progressively closer and weaker until they merge. The spectrum is shown in Plate IIa. The members of the series are respectively named Ha, Ha, Hy, Ha, etc., and as many as thirty-three lines have been measured. The position of a line is indicated by giving its wavelength λ in Ångström units (1 A=1×10⁻⁸ cm.), but it is often more convenient from the point of view of calculation to use the reciprocal of the wavelength which is called the wave number ν' . Thus $1/\lambda = \nu'$ and as ν' is a reciprocal of a length its units are cm.-1 For example, the wavelength of Ha can be converted into wave numbers as follows:

$$\lambda = 6562.8 \times 10^{-8}$$
 cm., hence $\nu' = 1/\lambda = 15.253$ cm.⁻¹

Although the hydrogen series spectrum is relatively very simple no relationship between the wavelengths of the lines was found until in 1885 Balmer discovered that they fit the formula

$$1/\lambda = \nu' = \mathbb{R}(1/2^2 - 1/n^2)$$

where R, a constant now called the Rydberg constant, is equal to 109,677 cm.⁻¹ The respective wavelengths of the different lines H_{α} , H_{β} , H_{γ} , H_{δ} , etc., are obtained from this formula when n is given the values 3, 4, 5, 6, etc. From the nature of the relationship, Balmer and Ritz suspected that other

series might exist wherein the first term is different from $1/2^2$. These have in fact been found, the general expression for them being $\nu' = R(1/m^2 - 1/n^2)$. Since R is known, the positions of the lines can be exactly predicted. Lyman found the first series, with m=1 and n=2, 3, 4, etc., lying in the deep ultraviolet region of the spectrum. The second series, with m=2 and n=3, 4, 5, etc., is the Balmer series. Paschen discovered a third series in the infra-red with m=3 and n=4, 5, 6, etc., and Brackett reported the existence of the series m=4 and n=5, 6, 7, etc., in the far infra-red.

Series similar to those in hydrogen have been found in the line spectra of other atoms, only in these cases the lines are often mixed up and must be sorted out by various means. In all of these series the wave number of a line can be represented as the difference between two terms. If c is the velocity of light, ν the frequency of light of wavelength λ , we have, by definition, $c=\nu\lambda$ so that $\nu'=\nu/c$. From this it follows that the frequency of a line is also the difference between two terms. The wave number in the Balmer series is given by the difference between terms of the form R/m^2 and R/n^2 . This applies to all series and is known as the Ritz combination principle. It gave Bohr the clue enabling him to explain atomic spectra in terms of the quantum theory.

Difficulties in the classical concept of the hydrogen spectrum

Rutherford's experiments with radioactive particles led him to the view that atoms consist of a small massive positively charged nucleus surrounded by outer electrons, both nucleus and electrons being small compared with the diameter of the atom as a whole. The charge on the nucleus of different atoms is always an integral multiple. Ze, of the electronic charge, e. Van den Broek suggested that Z might be equal to the atomic number, which is the numerical position which the atom occupies in the Periodic Table of the elements, and this was proved to be the case by Moseley's measurements on X-ray spectra. The simplest of all atoms is hydrogen which can be considered to be built up of a relatively massive nucleus, the proton, around which revolves a single electron.

According to classical mechanics, if a single electron revolves

in a circle round a proton as nucleus, the inverse square law of charge attraction being obeyed, then energy must be radiated by the electron and the orbital radius must change during the radiation. As a result of this the energy will be radiated in the form of a continuum and not as the discrete spectrum lines which are actually observed. Furthermore, final equilibrium would only be reached when the electron had fallen into the nucleus. This classical view fails, therefore, to account for the existence of a line spectrum and according to it atoms should be unstable, which we know not to be the case. It must be concluded that the electron is not free to choose any orbit but can only move in certain privileged orbits. These orbits can only be defined by quantum conditions, the electron being said to occupy quantised orbits.

Bohr's theory of the hydrogen spectrum

In order to account for the hydrogen spectrum on a quantum basis Bohr propounded two fundamental postulates:

- (1) Stationary states exist in the atom. Amongst the infinite number of orbits that classical theory expects, a selected number obeying specified conditions only can be occupied. The electron can rotate only in these stationary states and only in such a condition can the atom remain an infinite time without radiating.
- (2) The electron can move by a quantum jump from one stationary state or orbit to another, radiating out a quantum of radiation in the transition. If the energy of the atom in the first stationary state is E_2 and that in the second is E_1 , the energy radiated is $E=E_2-E_1$, and, just as in the case of Planck's oscillators, this will be equal to $h\nu$ where ν is the frequency of the radiated energy. Hence $E_2-E_1=h\nu$. This is called Bohr's frequency condition.

It will be observed that ν is not the frequency of orbital rotation in either of the stationary states. Converting the radiation frequency into wave numbers gives

$$\nu' = \nu/c = \mathbf{E}_2/hc - \mathbf{E}_1/hc$$

Thus, as in the empirical Balmer series formula, the wave number is expressed as the difference between two terms.

Consider the application of Bohr's postulates to the simple

case of a massive nucleus of charge Ze (in hydrogen Z=1) and mass M, assumed to be at rest. Around this nucleus an electron of mass m and charge e rotates in a circle of radius r with velocity v. For equilibrium the electrostatic attraction must provide the centripetal force on the electron so that

$$Ze.e/r^2 = mv^2/r$$
 (11)

$$r = Ze^2/mv^2 \quad . \quad . \quad . \quad . \quad (12)$$

If we consider the potential energy of the atom to be zero when the electron is removed to infinity, the potential energy of the system considered above must be equal to $-Ze^2/r$, for we must do an amount of work equal to Ze^2/r to remove the electron to infinity. The total energy of the system is E, the sum of the potential and kinetic energies, *i.e.*

$$E = mv^2/2 - Ze^2/r$$
 . . . (13)

Substituting from equation (11), we get

$$E = -Ze^2/2r = -mv^2/2$$
 . . . (14)

At this point Bohr introduces a further entirely arbitrary assumption, the only justification for which is that it gives the correct results. In the later discussion on the wave mechanics it will be found that all these arbitrary assumptions can be proved to have a sound theoretical basis. The assumption made here is that in those orbits which are possible it is the angular momentum which is quantised, the possible values being multiples of $h/2\pi$. The assumption can be expressed as

$$mvr = nh/2\pi$$
 (15)

n can have the values 1, 2, 3, etc., and is called the principal quantum number of the electron. Dividing equation (11) by (15) gives

$$v = \mathbb{Z}e^2 2\pi/nh \quad . \quad . \quad . \quad . \quad (16)$$

and as from (14) the total energy for any orbit is equal to $-mv^2/2$, the energy of the nth orbit is

$$\mathbf{E} = -2\pi^2 \mathbf{Z}^2 e^4 m/n^2 h^2$$
 . . . (17)

Thus for two stationary states E_1 and E_2 with n_1 and n_2 as principal quantum numbers we have the radiated transition energy

$$\mathbf{E}_2\!-\!\mathbf{E}_1\!=\!\!h\nu\!=\!\{2\pi^2\mathbf{Z}^2e^4m/h^2\}\{1/\!n_1{}^2\!-\!1/\!n_2{}^2\}$$

hence, converting into wave numbers, ν' , $(\nu' = \nu/c)$ gives

$$\nu'\!=\!\{2\pi^2Z^2e^4m/h^3c\}\{1/n_1{}^2\!-\!1/n_2{}^2\}.$$

This is Balmer's formula for the hydrogen spectrum, the Rydberg constant being equal to $\{2\pi^2Z^2e^4m/h^3c\}$. When this was evaluated in terms of the constants comprising it, all of which were known, the number obtained by Bohr was 109.737 cm^{-1} which is remarkably close to the value derived

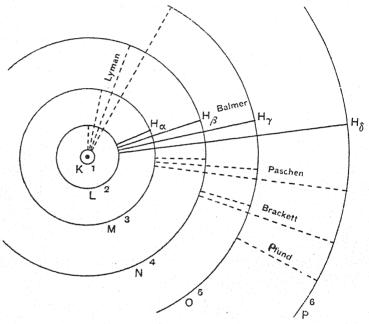


Fig. 5.5

from the Balmer series, 109,677 cm.⁻¹. This agreement constitutes a brilliant confirmation of Bohr's quantum theory of the origin of spectra.

Going back to equation (12) we find that the radius of an orbit r is given by $r=Ze^2/mv^2$, and as v is given by equation (16) we get

$$r = n^2(h^2/4\pi^2 Z e^2 m)$$
 (17)

The term in the brackets is constant, hence the radii of the possible orbits that the electron can fill are proportional to n^2 . This is illustrated in Fig. 5.5 which also shows how the

different series arise. The electron can be raised to a high level by some form of excitation and then falls back, in the Balmer series to the level with n=2. The diagram shows how the Lyman, Paschen, and Brackett series arise. The orbits corresponding to n=1, 2, 3, 4, etc., are called K, L, M, N, etc., orbits.

The orbital radii can be calculated from equation (17), and on substituting the values of the constants the radius found

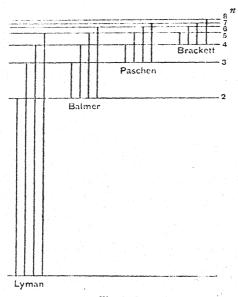


Fig. 5.6

for n=1 is 0.528 A, which is of the same order as the atomic radius of hydrogen given by the kinetic theory of gases.

The different series can also be illustrated diagrammatically by plotting the energies of the different orbits horizontally, line transitions being shown by a vertical line (Fig. 5.6).

The wave mechanics

The general principles used by Bohr have been successfully extended to include the very much more complex spectra emitted by atoms with many outer electrons. In such atoms the interactions between the electrons complicate the analysis considerably. In spite of this a high degree of success has

been achieved, the main features of most spectra being accounted for on the basis of electron orbits obeying Bohr's quantum conditions. There are, however, certain discrepancies which occur even in the relatively simple case of helium, which has only two outer electrons. The proposed quantum conditions are also very arbitrary assumptions the meaning of which it is difficult to understand from first principles. These difficulties have been resolved by the development of new methods of approach, the "wave mechanics" and the "quantum mechanics," both of which turn out to be different mathematical forms of the same theory.

The fundamental principle of the wave mechanics, first proposed by de Broglie, states that a moving electron or particle has associated with it a wave motion. The wave motion has a definite wavelength λ given by $\lambda = h/mv$ where h is Planck's constant and mv is the momentum of the particle. The properties of an electron considered to be rotating in an orbit must therefore be studied by examining the distribution of the associated waves. What was previously considered an orbital problem in mechanics now becomes a problem in wave motion. Just as in optics the simple ray theory must be replaced by wave theory to account for phenomena like diffraction, polarisation, etc., so in atomic phenomena wave mechanics replaces the orbital theory in order to explain many details of spectra, nuclear properties, electron diffraction, and reflection, etc. The details of the wave mechanics theory will be considered in a later chapter, however. The more mechanical picture of electron orbits still retains its value as a practical working model of the atom. It is generally used because of its relative simplicity, but the modifications demanded by the more accurate wave equations must be added to the deductions from the orbital model

REFERENCES

[&]quot;Heat and Thermodynamics." J. K. Roberts. (1933.)
"Atoms, Molecules and Quanta." A. E. Ruark and H. C. Urey. (1930.)

CHAPTER 6

THE THERMIONIC EFFECT

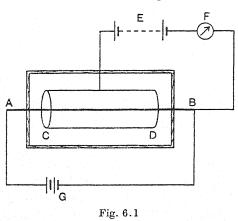
Introduction

It has been known since the eighteenth century that the air in the neighbourhood of an incandescent metal is an electric conductor. That this is due to the emission of ions by the hot body was proved by Elster and Geitel towards the end of the nineteenth century. The phenomenon is described as the thermionic effect, the metal being said to emit thermions. was noticed that the thermionic emission depended upon the temperature of the emitter, the nature and pressure of the surrounding gas, and the nature and surface condition of the metal. The observed phenomena become therefore very complex unless some of the many variables are eliminated. Our present day knowledge of this technically important branch of atomic physics is largely due to the pioneer work of O. W. Richardson who first obtained ordered results by paying special attention to purity and by making observations with very high vacua.

The thermionic current

The thermionic emission can be studied by means of very simple apparatus, similar to that shown in Fig. 6.1. The wire AB is heated by means of an electric current from G, and acts as the emitter. A metal cylinder CD surrounds the hot wire, and the battery E enables a potential difference to be applied between the two electrodes. The apparatus is baked out and highly evacuated, the wire being kept incandescent for over a week in order to remove traces of impurity. If the cylinder is made positive relative to the wire a thermionic current flows across the space between the electrodes and can be detected by the measuring instrument F. If the potential is reversed, the current ceases. It is clear, therefore, that the thermions are negatively charged.

It can easily be shown that the negative ions constituting the thermionic current are electrons given off by the hot metal. When a potential difference is applied between the hot cathode and the outer cylinder the ions are accelerated towards the latter. If now a magnetic field H is applied at right angles to the electrostatic field, and parallel to the wire, the ions will be deflected into curved paths. For a certain field strength they will just fail to reach the outer electrode and no current will pass. If the wire is thin, this occurs when $H^2=8Vm/d^2e$, where V is the applied potential, d the distance between wire and cylinder, and e and m the charge and mass of the ions.



By this means e/m can be measured and it is found to be identical with that of the electron.

Effect of temperature upon thermionic emission

If the temperature of the wire (deduced from its resistance) is maintained constant, the thermionic current varies when the applied voltage is altered. The current does not, however, obey Ohm's law, rising at first with the increasing potential difference but quickly reaching a maximum saturation value beyond which it does not increase. The potential required to produce saturation increases with rise in temperature of the cathode. To simplify the study of the effect of temperature upon the thermionic emission, sufficient voltage is applied to produce saturation for any particular filament temperature.

When the temperature of the cathode is raised there is a rapid increase in saturation current. The type of curve obtained when the saturation current I is plotted against the temperature of the emitter T is shown in Fig. 6.2, from which it is seen that I and T are related to each other in a regular manner. Richardson at first proposed the formula $I = AT^{\frac{1}{2}}exp(-b/T)$, A and b being constants characteristic of the metal constituting the cathode.

Richardson's formula not only fits the experimental curve,

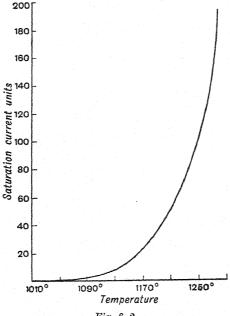


Fig. 6.2

but can be derived from simple kinetic considerations if it is assumed that a metal contains a cloud of free electrons in it behaving like a perfect gas. Such electrons have been postulated to explain the electrical conductivities of metals, the Peltier effect, metallic optical properties, etc. In the simple theory the electrons are considered to have their velocities distributed according to Maxwell's law. Thus even at ordinary temperatures the electrons will possess appreciable velocities. They are, however, confined within the metal, for when an electron passes through the surface it induces an equal and opposite charge upon the metal which tends to prevent it from escaping. An electron can only escape if it possesses more than a certain limiting velocity in a direction normal to the surface, the limiting amount corresponding to the energy required to overcome the work of emission against the induced charges.

As the temperature of the metal is raised the mean kinetic energy of the electrons will increase until a large number have energies exceeding the limit value and so succeed in escaping. If these are allowed to accumulate in the space above the metal an equilibrium will be established analogous to the production of a constant vapour pressure over a liquid in an enclosure. Under equilibrium conditions the number of electrons leaving the metal equals the number returning from the electron cloud above it. The latter can be calculated directly from gas kinetic considerations and leads to the formula

$$I = AT^{\frac{1}{2}}exp(-b/T).$$

However, a great deal of evidence shows that the electrons within a metal do *not* behave as a perfect gas. Richardson proved that if the electron emission is considered to be equivalent to the evaporation of a monatomic gas, the application of thermodynamic principles leads to the formula

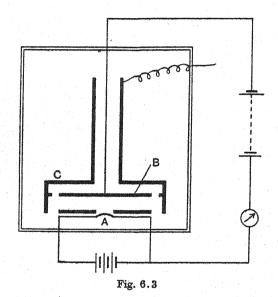
$$I = A'T^2 exp(-b'/T)$$
.

Owing to the fact that the exponential factor has a much greater influence than the other term depending upon T, both this and the earlier formula can be made to fit the observations. However, the modern quantum electron theory of metals shows that the latter formula is correct.

Energy of the emitted electrons

Since the electron densities in the space outside the metal are small in thermionic emission, it is to be expected that in this region the electrons will behave like a perfect gas for their mutual repulsion effects will then be negligible. This will be true even if the distribution within the metal itself is non-Maxwellian. Richardson measured the velocity distribution

amongst the emitted electrons with the apparatus shown in Fig. 6.3. The number of electrons given off by a heated platinum foil A is measured by the collecting plate B which is connected to an electroscope. A retarding potential V is maintained between A and B. C acts as a guard ring and electrostatic shield. In order to eliminate any small voltage effects due to contact potentials the plate B is covered with platinum. An electron will only reach the plate B if its component of velocity perpendicular to the plate, u, is such that $\frac{1}{2}mu^2$ exceeds Ve. In the experiment the retarding voltage is



varied and the current I between the electrodes is measured. If the velocities in the direction considered obey Maxwell's distribution law, it can be proved that $I = I_0 exp(-Ve/kT)$ where I_0 is the current passing when V is zero. This equation was found to hold exactly. Richardson modified the experimental arrangements enabling him to measure the velocities in the direction parallel to the plate and these also exhibited the same type of velocity distribution.

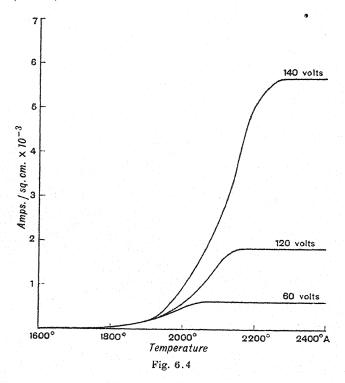
The results of the measurements of the velocities in the two directions prove that the general velocity distribution is Maxwellian, a conclusion that has been confirmed by later observation. It appears that the energy distribution is that corresponding to an electron gas whose temperature equals that of the hot surface. It is clear that electron emission must result in a cooling of the emitter as in the corresponding case of liquid evaporation. Conversely the absorption of thermions by a metal should raise its temperature. Both these effects have been observed.

The effect of the potential difference upon the thermionic current

It would at first appear that a small voltage should suffice to produce the saturation current, but owing to the mutual repulsions of the electrons in the space above the hot metal this is not so. If the emission is very large the accumulated space charge produced by this becomes considerable, and the true saturation corresponding to the temperature used can only be reached if very high voltages are applied in order to overcome the space charge effect. The space charge was first studied by Langmuir, typical curves being shown in Fig. 6.4. The curves each represent the relation between current and temperature when the potential difference between the hot wire and anode is maintained at a given constant value. At the lower temperatures the curves coincide with the saturation curve, but as the temperature is increased a smaller and smaller fraction of saturation is achieved. The bigger the applied voltage, the more rapidly are the accumulated electrons removed from the field and the higher the final value of the current. In each case a constant current is reached depending only upon the potential difference. If I is the value of the constant current corresponding to the applied voltage E it is found that $I = kE^{\frac{3}{2}}$, k being a constant. This relationship. known as Langmuir's law, can be derived by the following treatment given by Loeb.

Consider a pair of parallel plates d cms. apart with a potential difference E volts between them. The negative plate is hot, acting as a thermionic emitter of electrons, resulting in the passage of an electron current between the plates. Let the electron current density be I amp/cm.² If v is the velocity of the electrons at any point, the current density there will be

 $I=\rho v$, where ρ is the volume space charge density and clearly equals Ne if there are N electrons per cubic centimetre, of charge e. The electron velocity v is that given by the potential V through which the electron has fallen, such that $Ve=\frac{1}{2}mv^2$, m being the electron mass. Thus $v=\sqrt{2Ve/m}$ and $\rho=I/\sqrt{2Ve/m}$.



Since Poisson's equation holds for the space between the plates then we have numerically

$$\frac{d^2V}{dx^2} = 4\pi\rho = 4\pi I \sqrt{m/2Ve}.$$

Multiplying both sides by 2dV/dx gives

$$\frac{2d\mathbf{V}}{dx} \left(\frac{d^2 \mathbf{V}}{dx^2} \right) = 8\pi \mathbf{I} \sqrt{\frac{m}{2e}} \frac{d\mathbf{V}}{\mathbf{V}^{\frac{1}{2}} dx}$$

This is now integrated between the limits $(dV/dx)_0$ at x=0 where V=0 and dV/dx at x=x where V=V giving

$$\left[\left(\frac{d\mathbf{V}}{dx}\right)^{2}\right]_{(d\mathbf{V}/dx)_{0}}^{d\mathbf{V}/dx} = 8\pi\mathbf{I}\sqrt{\frac{2m}{e}}\left[\mathbf{V}^{\frac{1}{2}}\right]_{0}^{\mathbf{V}}$$

The potential V at x=0 is 0 and if we assume that the space-charge and electron current increase until the electric field at the cathode is zero, then $(dV/dx)_0=0$, leading to

$$\left(\frac{d\mathbf{V}}{dx}\right)^2 = 8\pi \mathbf{I} \sqrt{\frac{2m}{e}} \, \mathbf{V}^{\frac{1}{4}}$$

$$\frac{d\mathbf{V}}{\mathbf{V}^{\frac{1}{2}}} = \sqrt{8\pi \mathbf{I} \sqrt{2m/e}} \, . \, dx$$

Thus

Integrating from V=0 at x=0 to V=E at x=d gives finally

$$I = \frac{1}{9\pi} \sqrt{\frac{2e}{m}} \frac{E^{\frac{3}{2}}}{d^2}$$

Incidentally from this expression e/m can be evaluated.

Effects of gases and of impurities

The thermionic emission can be profoundly altered by the presence of a gaseous atmosphere around the emitter. Either adsorption phenomena intervene or else ionisation by collision can set in. Very small quantities of gas can have great effects, a trace of hydrogen being capable of making the thermionic current 105 times as great. The surface adsorption effects will be discussed separately later. Apart from the normal thermionic emission of electrons, a fresh wire always gives off positive ions for a certain period, the number of these diminishing rapidly with prolonged heating in vacuo. In order to examine the nature of these positive ions Richardson measured the values of E/M by employing the magnetic deflection method. He found that the ions given off initially by a platinum wire are positively charged potassium atoms occluded as impurity. After heating for 24 hours these were replaced by sodium ions. It may appear surprising at first that so called "pure platinum" should contain potassium and sodium as impurities, but when the total positive ion emission is integrated its weight is found to be only 1×10-5 per cent. of

the weight of the platinum. Such minute impurities cannot be removed by chemical means, and, as it is to be expected, the positive emission varies with different samples of the same metal. It appears at a much lower temperature than that required for electron emission but, as with the electrons, the velocities are distributed according to Maxwell's law.

Surface adsorption and the oxide coated cathode

A development of technical importance, particularly connected with the manufacture of thermionic valves and oscillographs, has arisen from the early discovery by Wehnelt that the emission of electrons from a metal surface can be greatly increased by coating it with certain oxides. Materials upon the surface of a thermionic emitter can act in two ways. In the first and more obvious effect, the adsorbed material alters the work function, that is to say, it alters the amount of work that one of the conduction electrons in the metal must do in order to escape. The adsorbed atoms form an electrical double layer upon the metal surface which assists or retards the emission electrons, according to circumstances. With adsorbed electropositive elements the positive side of the double layer is outwards so that the work of emission for an electron is reduced. Conversely the adsorption of electronegative elements reduces the thermionic emission by increasing the surface work function. Effects due to these double layers are found when pure metals like cæsium or barium are adsorbed on platinum or tungsten surfaces.

In the second type of surface effect it is not the conduction electrons of the supporting metal that are emitted, but the valence electrons of the adsorbed surface material. In such cases a good thermionic emission can be obtained with temperatures lower even than 200° C., whilst with pure metals nothing can be detected below 1,000° C. The oxide-coated cathode is a typical example of this type of emitter. It consists of a platinum wire coated with one or more oxides of an alkaline earth metal, such as BaO. This is prepared by immersing the wire in barium nitrate solution and then heating. The wire requires to be "activated" by the inclusion of a small quantity of carbon in the nitrate layer. The "activation" consists in reducing about ½ per cent. of the BaO to barium metal. The

metallic barium is adsorbed in the oxide layer and is the responsible agent for the increased emission. When the wire is heated to about 200° C, the free barium atoms lose electrons by thermal ionisation, a process quite different from the thermionic emission of the conduction electrons in a metal. Electrons then leak back into the adsorbed layer from the metal supporting wire through the semi-conducting oxide and are captured by the ionised barium atoms, which are again ready for emission. This leaking back of electrons through the semi-conducting BaO layer explains why the emissivity of such a wire can remain unchanged for over a hundred thousand hours of continuous heating, the metal wire base being effectively an inexhaustible store of electrons.

The activity of an oxide-coated filament is destroyed if the free barium atoms in the layer are chemically combined by any means. This is called "poisoning." As is only to be expected, traces of oxygen or water vapour can easily destroy the activation. Re-activation of a poisoned surface can be produced by reduction processes, proving that free barium metal is the cause of the enhanced emission. The well-known dull emitter thermionic valve contains an activated oxide-coated filament.

The modern electron theory of metals

In the early classical gas kinetic theory of electrons in metals it was implicitly assumed that the energy in the metal was distributed between the metal atoms and the electrons according to the equipartition principle. If this were so the free conduction electrons must contribute to the specific heat, which means that the atomic heats of metals should be higher than those of non-conductors. This is not the case, and to overcome this difficulty a quantum theory of the electrons in metals has been developed largely due to the initial work of Sommerfeld. This theory has successfully accounted for electrical conduction phenomena and electron emission effects.

According to the quantum theory of atomic structure an electron in an atomic orbit is defined by four quantum numbers. The electrons within an atom obey a law, the Pauli exclusion principle, which states that no two electrons can have identical quantum numbers. Fermi and Dirac have generalised this principle so as to include the free conduction electrons within

a mass of metal. This ultimately leads to a new type of energy distribution law, the electrons being said to obey the Fermi-Dirac statistics instead of the Maxwellian distribution. This quantum distribution is such that the contributions of the electrons to the specific heat is very small and proportional to the absolute temperature. As distinct from the classical case, there is an energy distribution even at the absolute zero, the electron energies varying from zero to a value

$$W_i = (h^2/8m)(3n/\pi)^{2/3}$$

h being Planck's constant, m the electron mass, and n the number of free electrons in a cubic centimetre of the metal. The variation in n results in W_i having values for different metals between 2 e.v. and 10 e.v.

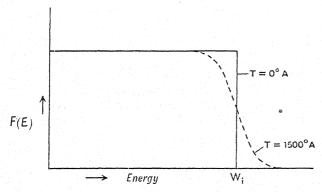


Fig. 6.5

The distribution function for the kinetic energies of the electrons can be discussed with the aid of Fig. 6.5. This diagram gives F(E) as a function of E, the kinetic energy, for temperature zero (solid line) and $1,500^{\circ}$ A (dotted line), in a representative case. At either temperature the number of electrons per unit volume of the metal having energies between E and E+dE is given by $E^{\frac{1}{2}}F(E)dE$. At absolute zero $F(E)=8\sqrt{2\pi m^{\frac{3}{2}}/h^3}$ over the energy range 0 to W_i and is zero for higher energies. At higher temperatures it will be seen that only a small fraction of the electrons have energy greater than W_i .

It should be noticed that E is the kinetic energy of the electron in the metal. At the absolute zero no electron can escape from the surface since even with kinetic energy W_i it is unable to overcome the surface forces. But, as the tempera-

ture is increased, a finite but very small fraction of the electrons possess this energy, until eventually a measurable thermionic emission, increasing rapidly with further increase of temperature, is obtained.

If W_a is the work to be done by an electron in escaping from a position of rest within the metal, those with kinetic energy W_i will only require to be given an amount of energy equal to $W_a - W_i$ in order that they should be emitted. From the distribution curve in Fig. 6.5 it can be shown that the electron current leaving unit area is

$$I = (4\pi k^2 me/h^3) \cdot T^2 \cdot exp\{-(W_a - W_i)/kT\}$$

This can be rewritten as $I=A'T^2exp.(-b/T)$ which is identical in form with the expression derived by Richardson from purely thermodynamic considerations. The quantity $(W_a-W_i)/e$, called the thermionic work function ϕ_0 , is a measure of the work required to remove an electron from the surface of the metal.

It can be shown by wave mechanics methods that not all the electrons having sufficient energy to escape actually do so. A certain number are "reflected" at the boundary and return into the body of the metal. If R is the mean effective "reflection coefficient" the emission formula becomes

$$I = A'(1-R)T^2 exp.(-b/T) = A''T^2 exp.(-b/T).$$

The value of A" depends in a sensitive way upon the chemical purity of the surface, R being affected in a critical manner by the potential distribution on either side of the boundary. This explains why the emission depends so much upon the state of the surface. In platinum, for instance, A" can vary from 1.45×10^7 to 1.07×10^{-3} .

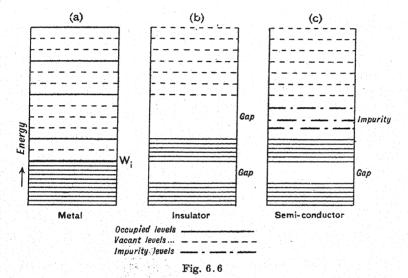
Electronic states in solids

In solids the valency electrons are considered to occupy a number of discrete states of different energy? The number of such states depends upon the total number of electrons in the material. In a metal at the absolute zero all the available energy states up to a certain value, W_i , are occupied. At any temperature above this a small fraction of the total number of electrons moves up to higher possible, but unoccupied states. The main concentration of electrons remains below the level

 W_i as shown by the quantum mechanical distribution curve in Fig. 6.5.

A schematic diagram giving a rough model of the distributions of electrons in the states of a metal, an insulator, and a semi-conductor is shown in Fig. 6.6. In the case of a metal at a temperature above the absolute zero, there is a large concentration below W_i (shown by continuous lines) and a small concentration above this. The broken lines represent possible energy states as yet unoccupied.

The distribution in the case of an insulator (diagram b) is distinctly different. Energy levels are grouped closely together



but the groups are separated by wide gaps which are completely devoid of electron levels. At the absolute zero of temperature the occupied levels cease at a gap. When the temperature is raised the upper levels still remain empty, at moderate temperatures, since the gaps cannot be crossed. At very high temperatures a number of electrons acquire sufficient energy to enable them to cross gaps and enter upper levels.

A group of solids important to thermionics and photoelectricity is the *semi-conductor* class. In this, the distribution of electron energy levels is similar to that in insulators but semi-conducting properties are exhibited by virtue of the existence of *impurity atoms*. These lead to occupied electronic states in the gap between the occupied and upper possible levels and act therefore as a bridge. These impurity levels are represented by the broken and dotted lines (diagram c). The impurity levels do not extend continuously throughout the material but are *localised* near the impurity atoms. At moderately high temperatures electrons can move from the impurity levels into the higher dotted levels normally unoccupied at low temperatures.

Electrical conductivity

As each electronic state is horizontal, in one of the above energy diagrams representing the levels in a slab of material, an electron should be able to move along an energy state without doing work. This means that the material will be a perfect conductor of electricity, exhibiting no resistance to the movements of electrons. However, the thermal kinetic activity of the molecules comprising the material has up to now not been taken into consideration. As an electron moves along a horizontal line representing a state, it encounters potential barriers due to the agitated molecules and unless it is given energy it cannot rise over the barriers.

The application of an electrical potential to a slab of metal, say between the right and left-hand sides, as in Fig. 6.7, results in giving the energy levels a gradient in this direction. They are no longer horizontal and the displacement of an electron from an occupied level to an empty level, previously higher, is now possible and takes place. There is no change in the total energy of the electron which moves, but the kinetic energy increases with corresponding diminution in potential energy. The electrons can now move towards the positive electrode and we have electrical conduction. Work has still to be done in overcoming potential barriers. The energy lost in this manner represents the heating normally ascribed to electrical resistance.

Below the energy level of value W_i there is practically no displacement of electrons, since in a metal most of these energy states are occupied. The electronic conduction, therefore, mostly takes place in the upper band of unoccupied levels which acts as a conduction channel.

When a potential difference is applied to an insulator the

movement of electrons into unoccupied states does not take place, since the gap is such that it cannot be jumped.

In a semi-conductor there are, it is true, electronic states in the gap, but as these are discontinuous in nature conduction cannot take place at very low temperatures. At somewhat higher temperatures electrons from the impurity states are able to pass into the upper continuous levels and then conductivity sets in, as in the case of a metal. There is one striking distinction between the conducting properties of a

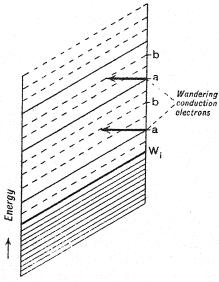


Fig. 6.7

metal and a semi-conductor. In a metal the electrical resistance increases with rise in temperature as the more active thermal motions of the molecules interfere more strongly with the motion of the electrons. In a semi-conductor the electrical resistance diminishes with rise in temperature, since more and more electrons succeed in entering the upper states and this effect considerably outweighs any effect due to increased thermal agitation.

Electrical image forces

At first it would seem as if the application of only a very small potential should suffice to draw all electrons out of a metal.

This would indeed be the case if there were no forces tending to retain the electrons within the metal boundary. There are such forces, the chief of these being the electrical image force. When an electron is withdrawn from a smooth-surfaced metal, it induces an equal and opposite charge in the metal and the resulting attractive force tends to oppose the withdrawal of the electron. If the electron, charge e, is withdrawn a distance x from the smooth surface the force is such as if an equal and opposite charge exists a distance x behind the metal surface. It is therefore equal to $e^2/4x^2$. On moving the electron through a distance dx the work done is

$$\delta W = e^2 dx/4x^2.$$

Thus the work required to move the electron from a point at a distance r to infinity equals

$$W = \int_{r}^{\infty} \frac{e^2}{4x^2} dx = \frac{e^2}{4r}.$$

Clearly as r becomes very small, i.e. as the surface is approached, W tends to become infinite. The physical interpretation of this would be that there could never be thermionic emission. Actually, however, the assumption that the metal surface is smooth cannot be true for distances of the order of atomic dimensions and the law of force within these distances cannot be expressed in simple form. But we can write $W = W_0 + W_1$ where Wo is the work needed to bring the electron out of the metal against some unknown law of force and W1 the work done in accordance with the inverse square law. Theory shows that the total work W does not depend in a critical manner upon the particular law of force operating at short ranges. the short range force is effective over a distance x_1 the total work is always approximately equal to $W=e^2/2x_1$ no matter which of the probable laws of force is adopted. The value of x_1 depends upon the particular metal considered. Also, in the notation of p. 106, $W = W_a$.

If ϕ is the potential necessary to extract the electron $W = e\phi$. Thus $\phi = W/e = e/2x_1$ enables the value of x_1 to be calculated. When observed values of ϕ (or W_a) are substituted, x_1 is found to be of the order 1 or 2×10^{-8} cms., which is of the order of the radius of an atom. The short range forces therefore operate over atomic dimensions only. As $\phi \propto 1/x_1$ electrons can

more easily escape from heavier atoms than lighter, for the atomic radius is larger and therefore ϕ is smaller.

If an electron is to escape by virtue of its own kinetic energy then the latter must be in excess of ϕe .

The Schottky effect

In direct thermionic emission an electron escapes when its thermal energy suffices to enable it to overcome the electrical image forces. If an external anode is placed close to a thermionic emitter, exerting a force which tends to draw electrons away from the hot metal, it is clear that at some distance from the hot surface the two fields balance. Any electron having sufficient thermal energy to enable it to reach this point will succeed in escaping. The greater the external field the nearer is this critical point to the hot surface. As this point moves in, less and less energetic electrons can reach it and thus escape. Effectively there is a reduction in work function because of the presence of the external field. An increase in the latter raises the thermionic emission.

This effect was first pointed out by Schottky who derived an expression for the effect which is $I=I_0exp.(cE^{\frac{1}{2}}/T)$ where I_0 is the thermionic current when there is no external field, I the current when the external field strength is E, T the absolute temperature, and c a constant, A field of some 2,000 volts per centimetre, which is easily attained, increases the zero field emission by 10 per cent.

Cold emission

Clearly the greater the field strength the more important becomes the Schottky effect. With a sufficiently intense field, however, an entirely new effect sets in. The maximum of potential near the surface of the metal becomes exceedingly sharp and a wave-mechanical potential-barrier effect occurs. The more energetic of the electrons constituting the main body, with kinetic energies less than W_i , and independent of the temperature, possess a finite probability of escape through the barrier. There is then cold cathode emission, sometimes called field emission. Both theory and observation show that the cold emission current is given by $I = aE^2 e^{-b/E}$ where E is the field strength, and a and b are constants for a given metal.

Fields exceeding 10⁵ volts per centimetre are usually required before cold emission becomes manifest. Up to about 700° C. the emission is independent of the temperature, but after this point true thermionic emission begins to assert itself.

If a thin surface layer of copper is separated from a massive metal copper base by a thin semi-conducting blocking layer of Cu₂O, only some 10-6 cm. thick, the application of but a few volts across this film gives a field strength sufficient to produce cold emission. When the surface of a copper plate is oxidised, oxygen atoms are occluded in the copper-oxide film as impurity atoms forming a true semi-conductor. If a piece of copper gauze be placed upon the top of the film to act as a contact maker, the combination acts as a rectifier to alternating currents. The application of even small voltages produces cold emission, but as many more electrons are available in one direction than in the other, owing to the massive metal base acting as a relatively infinite source of electrons, the flow is easier in this direction than in the other. The system therefore acts as a rectifier. This fact has been made use of in electrotechnics and in radio work where metallic rectifiers have become of importance. The rectifying actions of the early radio crystal detector and of the coherer are considered to be similar in nature to copper oxide rectification.

The shot effect

Interesting support for the correctness of the atomic conception of electricity, and indeed a means for measuring the charge on the electron, is afforded by the so-called "shot effect." Owing to the atomic nature of the thermionic emission the electron stream is not perfectly uniform. It is to be regarded as a hail of particles like small shot, and therefore exhibits statistical fluctuations in any given time interval. The smaller the time interval considered the larger are the possible fluctuations in the number of particles emitted in any one of these time intervals. These random fluctuations are equivalent to the superposition of a small alternating current on top of the average direct current thermionic emission. A statistical examination of the emission shows that the average fluctuation

of charge δ in a time interval t is given by $\delta = \sqrt{ie/t}$ where i is the mean direct current and e the electronic charge. It is possible by this means to obtain a value for the electronic charge.

The fluctuations producing the shot effect are of very great importance in thermionic valves using high amplifications. These fluctuations clearly set an upper limit to the degree of amplification that can ever be obtained by thermioric valves.

REFERENCES

[&]quot;Emission of Electricity from Hot Bodies." O. W. Richardson. (1922.)
"The Metallic State." W. Hume-Rothery. (1931.)
"Electron Emission and Absorption Phenomena." J. H. de Boer. (1935.) "Photocells." Campbell and Ritchie. (1934.)

CHAPTER 7

PHOTO-ELECTRICITY

Introduction

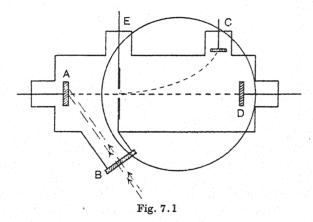
When substances, particularly metals, are irradiated with light of wavelength shorter than a certain limiting value. electrons are emitted. This, the photo-electric effect, has already been briefly mentioned in connection with the development of the quantum theory. The discovery of photoelectricity goes back to an observation made in 1887 by Hertz. who noticed that when ultra-violet light is allowed to fall upon a spark gap, the spark passes more easily than when the gap is not so illuminated. A year later Hallwachs discovered that a negatively charged metal loses its charge when irradiated with ultra-violet light, but the light has no effect at all on a positively charged body. A further step was made by Elster and Geitel, who proved that the electropositive metals potassium, sodium, and rubidium easily show photo effect when irradiated by ordinary white light. It is clear that ions must be emitted from bodies which lose charge when irradiated, and the Hallwachs effect proves that these must be negatively charged ions. In 1899 Lenard measured e/m for the photoions and by this means proved that they were electrons. Since the photo-electric effect is much more marked in metals than in other solids, liquids, or gases, the discussion following refers largely to metals.

Lenard's determination of e/m for photo-ions

The apparatus employed by Lenard for measuring the ratio of the mass to the charge for the ions emitted by a metal irradiated with ultra-violet light, is shown in Fig. 7.1. The apparatus is highly evacuated. Ultra-violet light, entering by the quartz window B, strikes an aluminium plate A. This is given a negative potential so that the negatively charge photoions which are liberated are repelled in a direction away from

A towards the earthed electrode E, which has a small circular aperture. This aperture isolates a beam of the charged particles which falls upon the electrode D and is detected with an electrometer. A magnetic field can be applied with its lines of force perpendicular to the plane of the paper (the circle diagrammatically indicates the position of the polepieces), and when this is done, the ions traverse a circular path reaching the electrode C which records them.

In considering the forces acting upon the ions it will be at first assumed that any velocity imparted to them by the actual act of emission can be neglected. This being so, the ions will acquire a velocity v due only to the potential P



applied between the electrodes A and E. If e and m are the charge and mass of the ions respectively, it follows that $Pe=mv^2/2$. In the magnetic field, of strength H, the particles describe the arc of a circle of radius R given by R=vm/He. The radius can be calculated by measuring the distances of the electrodes C and D from the aperture E. Combining the two expressions gives $e/m=2P/R^2H^2$ and v=2P/RH. experiment showed that e/m for the photo-ions was identical with that for the electron, thus it must be concluded that the photo-ions are electrons.

In Lenard's original experiment the applied voltage P was equal to 12,600 volts which, the above expression shows, gives

the electrons a velocity of 5.4×10^9 cms./sec.

It has been assumed that the initial velocity of ejection, if any, can be neglected and it is easy to show that with such a high applied voltage this is justified. The emission velocity can be measured directly by giving the electrode A a positive potential P_0 , which just suffices to cause the emitted electrons to return back to the electrode. Clearly $P_0e=mv_0^2/2$ where v_0 is the velocity acquired by the electrons due only to the act of photo-electric emission. In the particular experiment considered v_0 was found to be equal to 10^8 cms./sec. which is only one-fiftieth of the velocity produced by the accelerating potential P and can thus be neglected, unless an accurate value of e/m is required. In practice it is convenient to describe the energy of electrons in electron volts. An electron in falling through a potential difference of 1 volt acquires a velocity of 5.9×10^7 cms./sec.

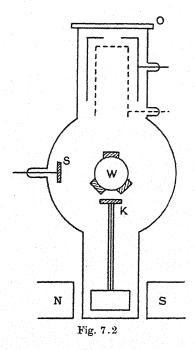
With the apparatus just described Lenard made a very important discovery. He found that the emission velocity of the photo-electrons could vary from zero up to a maximum value, v_0 . The value of this maximum was independent of the intensity of the light, being affected only by the wavelength used and by the nature of the electrode A. He showed further that the number of emitted electrons is proportional to the intensity of the illumination. These two fundamental laws were applied by Einstein to the development of the quantum theory. Being of such importance, a critical re-examination was undertaken by Millikan with the object of testing the validity of the conclusions arrived at by Lenard on the basis of only approximately exact measurements. Millikan's experimental refinements enabled him to achieve a high degree of accuracy and the laws were completely vindicated.

Millikan's investigation of the velocities of ejection of photoelectrons

As a first essential, Millikan considered it necessary to irradiate surfaces which were chemically clean, since by analogy with thermionic emission, it is to be expected that surface contamination will affect the work of emission and therefore the velocity of ejection of photo-electrons. In order to be able to study the effect of the wavelength of the incident light, the experiments were carried out with the alkali metals, since

they exhibit photo-electric emission with light from the visible region down to the ultra-violet. The use of these metals therefore gave an extended wavelength range over which observations could be made. The apparatus used was designed to enable the experimenter to cut clean surfaces of the metals in vacuo. It is illustrated in Fig. 7.2.

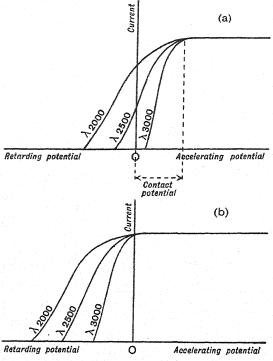
A wheel W, which can be rotated from outside, has attached to it three cylinders, each made of a different alkali metal.



K is an externally controlled knife mechanism by means of which the metal surfaces can be scraped clean. It is rotated electrically. The whole apparatus is highly evacuated in order to avoid secondary collision effects. A cleaned metal surface is brought opposite the window O and illuminated with monochromatic radiation of known wavelength. Photoelectrons are ejected. By means of a field applied between the emitting surface and a cylinder of gauze (dotted) the photoelectrons are retarded and can just be prevented from reaching the gauze which is made of oxidised copper, since this material

itself shows no photo-effect with the light used. The voltage between the emitting surface and the collecting gauze is varied regularly, the electron current which reaches the latter is measured with an electrometer. This is repeated in turn for each different wavelength of light used.

The results obtained for the different alkali metals are similar in nature, a typical voltage-current curve being that



Figs. 7.3a, 7.3b

shown in Fig. 7.3a. The potential difference between the emitter and the gauze can be made accelerating or retarding. It can be seen that a high accelerating potential gives a maximum current which represents the total photo-electric emission. As the voltage is diminished, a critical point is reached at which the photo-electric current for all wavelengths begins to fall, ultimately reaching zero. Now the potential applied between gauze and emitter is not the true potential

difference acting between them across the evacuated space, since there exists a contact potential difference between the respective materials. (This contact potential difference is that which supplies the driving voltage in primary electrolytic cells. It introduces a small potential difference, between the gauze and metal, of the order of 1 volt.) When the contact potential difference (measured against the electrode S) is corrected for, the current-voltage curves are displaced as shown in Fig. 7.3b. Thus when the gauze has a positive potential relative to the emitter a constant current passes, but when it is negative the current falls with increasing field, reaching zero at a point the value of which depends upon the wavelength of the light used (marked on the respective curves in Ångströms).

Einstein's photo-electric law

The intercept of each curve with the potential axis is the "stopping potential," the value of which gives directly the maximum velocity of emission. For clearly the gradual fall off in emission current with increasing retarding voltage proves that some electrons have a small initial emission velocity and are easily stopped. In fact the velocities vary from zero to that requiring a retardation given by the stopping potential. When these maximum voltages (the intercepts V) are plotted against the frequency of the light, they fall upon a straight line obeying the relation $V=k\nu-V_0$. V is the stopping potential for light of frequency ν , k a constant which has the same value for all metals, and Vo a quantity which differs from one metal to another. Electrons which require a stopping potential V must have an energy equal to Ve (e being the electronic charge), thus we have $Ve = ke\nu - V_0e = mv_m^2/2$. This relationship had been derived before Millikan undertook his investigations, and the earlier measurements had shown that the numerical value of ke was approximately equal to that of Planck's radiation constant, h. Einstein therefore proposed that the true relationship is $\frac{1}{2}mv_m^2 = h\nu - V_0e$. This assumption was shown by Millikan to be correct, since the value he found for ke was 6.57×10^{-27} erg/sec., which is identical numerically with Planck's constant.

The expression $mv_m^2/2 = h\nu - V_0e$ (Einstein's photo-electric law) is of fundamental significance in connection with the

quantum theory and with Bohr's theory of atomic spectra. For some frequency ν_0 , $h\nu_0$ will equal V_0e , in which case the emission velocity v will be zero. Thus it is clear that no emission will take place if the incident light has a frequency less than ν_0 . This critical light frequency is the photo-electric threshold frequency, the corresponding wavelength to it being described as the long wavelength limit, since light with longer wavelength is ineffective. It will be shown later that the threshold frequency is related to the thermionic work function and, like the latter, depends upon the nature of the emitter and the state of its surface.

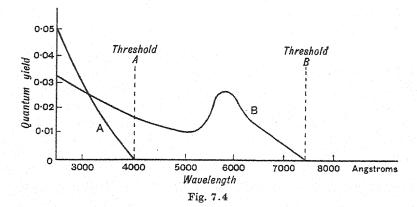
If the Einstein photo-electric law holds good, the intensity of the incident light should have no effect at all upon the emission velocity of the photo-electrons. This is indeed the case. Experiment shows that only the number of electrons emitted depends upon the intensity. The number is directly proportional to the intensity of the incident light, but for a given metal the velocities depend only upon the frequency. This has been tested over wide spectral regions with large intensity variations and is found to be true under all conditions. A more intense source of light merely supplies more quanta, not more energetic quanta. It is the energy associated with each quantum that determines the maximum emission velocity.

No matter how intense the light source may be, if the light frequency is less than ν_0 no emission takes place. If the surface of the emitter does not happen to be uniformly clean and is patchy, more than one value of ν_0 can be observed. In Millikan's experiments changes in ν_0 were observed if the surfaces were left exposed for some time to any traces of gaseous impurity left in the evacuated apparatus. Changes in the value of this frequency can also be brought about by recrystallisation effects upon the surfaces after they have been scraped clean in vacuo.

It should be noted that the velocity v_m which occurs in the Einstein formula is the maximum velocity that an emitted electron can have, the actual velocity can vary from zero up to this limit. The velocity distribution can be either derived from the shapes of the curves in Fig. 7.3b or measured directly by deflecting the photo-electrons with a magnetic field.

Effect of wavelength upon the number of photo-electrons emitted

Different wavelengths have different efficiencies in exciting emission. Thus two sources with the same absolute intensity but with different wavelengths produce different numbers of electrons. The effect of wavelength is clearly shown in Fig. 7.4. The curve A is that shown by a given pure metal with a chemically clean surface. The wavelength is plotted against the quantum yield, which is defined as the number of electrons per quantum of incident radiation. The quantum yield is clearly a measure of the excitation efficiency. The curve is smooth, the quantum yield increasing from zero at



the threshold (about 4,000 A.) as the wavelength of the incident light diminishes. There is evidence suggesting that the yield should reach a maximum value in the far ultra-violet region.

If the surface of the metal is contaminated either by adsorbed gas or impurities, or if a very thin film of some foreign metal is deposited upon it, the curve B is observed. The threshold in B has moved considerably to the red end, and instead of a smooth, regular curve showing steady increase in quantum yield with wavelength, the curve now exhibits a marked maximum. With some metals two maxima have been observed. Photo-electric emission in the neighbourhood of one of these maxima is called "selective" emission, elsewhere

it is called "normal." The names have an historical origin, since early workers considered that the maxima were due to some abnormal selective effect. When the emission is normal it is independent of the plane of polarisation of the incident radiation, but when it is selective the emission is usually greater when the light is polarised with the electric vector perpendicular to the surface. A great deal of the confusion existing in earlier investigations upon polarisation effects was caused by inattention to chemical cleanliness.

Apart from the "selective" humps, the regular increase in quantum yield with diminishing wavelength is quite understandable. The theory of metals shows that the electrons within a metal have different energies, hence light with $h\nu$ gradually increasing beyond the threshold value will be able to eject more and more electrons with less and less inherent energy.

The theory of the photo-electric effect

As in the case of thermionic emission, a metal which has an adsorbed surface film is able to emit two kinds of electrons, either conduction electrons from the hase metal or valence electrons from the surface film. When valence electrons are being emitted we get "selective emission" produced by the selective absorption of the incident radiation by the atoms in the surface film. This is a form of resonance, a particular wavelength being more effective than others in ionising the atoms in the surface film. In contradistinction, the "normal effect" represents the emission of conduction electrons which have been ejected from a pure metal without surface contamination.

When an electron of energy W, within a metal, absorbs a quantum of radiation of energy $h\nu$, its energy is increased to $W+h\nu$. When this energy equals the work W_a required to be done by an electron in overcoming the surface forces, the electron will emerge from the metal with zero velocity. In general the kinetic energy of a photo-electron will be $mv^2/2=W+h\nu-W_a$. According to the quantum theory of metals, electrons at absolute zero of temperature have energies varying from 0 to a value W_i . Thus at this temperature we have, for the electrons with maximum energy, $W=W_i$. From this it follows that $mv_m^2/2=h\nu+W_i-W_a$ where v_m is the

maximum velocity attained by an emitted electron. Einstein's photo-electric law shows that $mv_m^2/2 = h\nu - h\nu_0$, hence $h\nu_0 = W_a - W_i$. As the thermionic work function ϕ_0 is equal to $(W_a - W_i)/e$ then $h\nu_0 = e\phi_0$. This relationship shows that the photo-electric threshold energy at the absolute zero is equal to the work required for a thermion to overcome the surfaces forces. Experiments carried out with palladium have proved that this theoretical prediction is obeyed.

It will be noticed that there is a close analogy between the photo-effect and the thermionic emission when there is surface contamination. Just as adsorbed layers can increase thermionic emission, so they can also influence the position of the photo-electric threshold. Some metals can be made photosensitive in the infra-red by preheating in hydrogen, the resulting adsorbed surface layer producing a large shift in the threshold frequency value.

Photo-electric fatigue

An effect described as photo-electric "fatigue" was reported by the earlier investigators of photo-emission phenomena. It was observed that the emission from a pure metal began to fall off with increasing age of the specimen, and this was attributed to some form of "fatigue." It is now known that this so-called "fatigue" is due to slow oxidation of the surface. This results in an increase in the surface work function, with corresponding diminution in photo sensitivity. The importance of even slight surface impurities was not clearly realised by the earlier workers.

There is, however, a form of fatigue which is not due to surface oxidation and which sets in if the photo-emission takes place in the presence of gases. It is produced by temperature changes arising from the bombardment of the metal surface by atomic and molecular ions created in the gas by the photo-electrons. This fatigue effect (due to ionic bombardment) is not exhibited when the photo-emission is being given off by a pure metal surface, but only when the emission is from a metal covered with a surface film. The latter is very sensitive to ionic bombardment. The fatigue is due to an alteration in the state of the surface film caused by heating by ionic bombardment. In particular, metals which have been sensitised

with hydrogen are very liable to exhibit this form of fatigue. The surface heating does not materially affect pure metals. Theory predicts that photo-electric emission from pure metals should be almost independent of temperature, providing the temperature is not great enough to cause thermionic emission to set in. This has been verified by experiment.

Photo-electric effects with non-metals

Photo-electric emission is not confined to metals, being also exhibited by non-metals, gases and liquids. Some metallic compounds, such as oxides, sulphides, etc., are photo-active when dry, others require first to be moistened. Quite a number of organic compounds give an emission, either in the solid state or in solution. Certain liquids are also photosensitive. The photo-activity of gases has already been met with in the method of producing gaseous ions by the irradiation of a gas with X-rays or ultra-violet light. This is a direct photo-effect, frequently in an inner electron shell. The electrons which X-rays eject from solids in general are also photo-electrons from an inner shell.

The phenomenon of phosphorescence has a fundamentally photo-electric mechanism. A very large number of phosphorescent materials (phosphors) can now be made, according to varying recipes. A typical phosphor is as follows. Barium sulphide is mixed with sodium carbonate which acts as a flux, and to this a trace of an alkaline earth metal is added. Such a mixture is strongly phosphorescent and at the same time markedly photo-active. It appears that certain molecular groupings behave as active centres, electrically insulated from one another by the flux. On irradiation photo-electrons are emitted from the active centres. Most of these are trapped in the flux mass which is non-conducting. They tend to leak back to the ionised atoms in the active centres and when recombination takes place there is emission of light. A finite time is taken for all the electrons to leak back, so that the material continues to emit light for a considerable period. This is the phenomenon described as phosphorescence. At very low temperatures the phosphorescent light-emission is quenched, but when the temperature is increased the trapped electrons can collide more frequently with ionised atoms, so that the intensity of the phosphorescence increases.

Photo-chemical reactions are also examples of effects directly due to photo-electric emission. It has been proved that each quantum of light absorbed is able to produce a reaction with a molecule. This may set a chain of reactions in motion. A well-known typical photo-chemical reaction is the production of HCl by the irradiation of a mixture of hvdrogen and chlorine. All life is, moreover, dependent upon a photo-chemical reaction—the formation of sugars in plants from carbon dioxide by photo-chemical absorption. The action of a photographic plate is a practical application of a photo-chemical reaction. Silver halides are strongly photoactive, a photographic plate being essentially a suspension of a silver halide in a gelatine emulsion. The irradiation produces ionisation which forms the latent image. There is either a local chemical change where the light falls or else a physical change, and at each point affected a grain can be developed up. The latent image is not absolutely permanent, disappearing if and when the trapped electrons in the emulsion return to their parent atoms. This leaking back is similar to that which produces phosphorescence but takes place at a much slower rate.

The physiological effects of light, which are responsible for vision, etc., are usually photo-chemical in nature. It has been suggested that the act of vision is largely a simple photo-electric effect, since there is evidence of electrical changes in the retina when the eye receives light.

The inner and outer photo-electric effects

When a light quantum falls upon a metal it can only be absorbed by an electron if the energy in the quantum is sufficient to raise the electron to some higher vacant state. In a metal there are always vacant states immediately above the band of filled levels. Hence any light that is at all absorbed raises electrons into higher states with a resulting increase in electrical conductivity. At ordinary temperatures the increase in conductivity is only very slight since the conduction band already contains a number of electrons.

In the case of insulators the energy in the quantum must exceed that of the gap between the occupied lower and vacant upper bands before an electron can be lifted into the latter.

Similarly, in semi-conductors the energy in the incident quantum must exceed that required to raise an electron from an impurity level to a vacant upper band level. If the energy in the quantum does exceed this, electrons are raised into the conduction band producing a relatively large increase in the electrical conductivity. This is called the inner photo-effect. It is made use of in "conductivity cells," a selenium cell being typical. For if selenium is illuminated the inner photo-effect leads to a fall in the electrical resistance. This can readily be measured and may be used as a measure of the intensity of the incident radiation.

An electron cannot leave the boundary surface unless it has acquired energy sufficient to overcome the image force. Still more energy is therefore needed to detach the electron altogether, even though it has already been raised into the conductivity band. This emission of electrons, the normal photo-emission, is the outer photo-effect. It should be quite clear that the inner photo-effect does not lead to electron emission whilst the outer photo-effect does.

Both the inner and outer photo-effects have been made use of in practical applications to photo-electric cells. Such a cell is a device for the detection and measurement of light. It is clear that instruments capable of doing this will have wide applications. The measurement and recording of light intensities is technically of importance, particularly in talking films, television, etc. Photo-cells have many valuable purely scientific applications also. A number of distinctly different types have been evolved. We shall examine these according to the photo-electric properties involved in their operation.

Emission cells

These make use of the outer photo-effect and can be divided into vacuum cells and gas-filled cells. The vacuum cell consists of a prepared photo-sensitive surface upon which light is allowed to fall, leading to electron emission. The electron current between the sensitive surface and an anode in the cell is measured, a potential difference of about 100 volts being applied between the sensitive cathode and the anode. Many different types of sensitive surface have been employed. For the detection of white light a compound surface is used, a casium film being deposited upon a silver base which is covered

with cæsium oxide. The emission can be considered to be given off by a composite surface described as a Cs—O—Ag surface. For spectrophotometric work a K—O—Ag surface is found to be extremely good. If a narrow spectral region is under examination increased sensitivity can be achieved by the employment of a photo-sensitive material which has a maximum (selective effect) in the region studied.

Vacuum cells are extremely accurate in their response, the photo-electric emission being strictly proportional to the intensity of the light. If the cathode is suitably chosen the sensitivity of a cell does not alter over very long time periods. They are by no means as sensitive as cells which will be described in following paragraphs, but are essential if precise intensity

measurements are required.

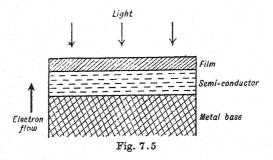
A tenfold increase in sensitivity results if a small quantity of gas is introduced into a vacuum cell. Such a detector is described as a gas-filled emission cell. The increase in sensitivity is due to the production of increased current by ionisation by collision. At a suitable gas pressure the photo-electrons ionise the gas by collision and the current increases. It is known that the time taken to eject a photo-electron from a surface is less than 10^{-8} second, and as the time taken for one of these electrons to reach and ionise an atom is usually less than 10^{-6} second with the cells in general employ, the time lag between the reception of the light and the production of the amplified photo-current is very small. The cell can therefore safely be used with light of very rapidly fluctuating intensity if need be.

The typical commercial cell, made by evaporating cæsium on to a silver surface which has upon it a deposit of cæsium oxide, has a high red sensitivity, the work of emission being low and the threshold being far in the red region of the spectrum. The increase in sensitivity produced by the introduction of the gas has made this type of cell commercially important. It has a good frequency response but has an accuracy of only 2 per cent. A constant light intensity can produce photo-currents which can fluctuate to within 2 per cent. because of the irregularities in the amplification in ionisation by collision. The small photo-currents are usually amplified further by means of thermionic valves.

Rectifier cells

The inner photo-effect has been applied to the detection of light and two forms have emerged, the rectifier cell and the conductivity cell, which will be discussed later. A photo-active boundary results when a semi-conducting layer is formed on the surface of a metal plate either by heat treatment or by cathodic sputtering. A typical case is the formation of a semi-conducting layer of Cu₂O by heating a copper plate in air. It is considered that oxygen acts as the impurity in the insulating oxide, converting it into a semi-conductor.

The rectifier cell with a sputtered film is more often used, as the film thickness and its nature are under control. In such a cell the light falls upon the *front wall* of the semiconductor in the manner shown in Fig. 7.5. The surface film



is effectively transparent and when light is absorbed at the interface of the film and semi-conductor electrons travel against the light. Illumination therefore results in the passage of a current through the cell. This current is not proportional to the intensity of the light flux in a simple linear manner but depends upon the resistance of the current detecting instrument in the circuit. The smaller the latter the more nearly linear is the relation between current and light intensity.

Rectifier cells require no external application of potential. They are robust and cheap and have high sensitivity, being able to detect light with the same degree of sensitivity as gasfilled emission cells. They have one drawback in so far as the current cannot easily be amplified by thermionic valves. The currents produced by strong light sources, such as daylight,

suffice to operate galvanometers directly without amplification. The cells have therefore been widely applied to the manufacture of instruments used in lamp photometry, general illumination meters, and exposure meters for photographic work, etc.

All rectifier cells exhibit a marked selective effect in the green or red. They are sensitive to temperature changes and have no effective time lag in operation.

Conductivity cells

Conductivity cells also employ the inner photo-effect but differ from rectifier cells regarding boundary conditions. In a thick slab of semi-conductor the electrons can accumulate at a boundary and set up a back electro-motive force which ultimately can prevent the flow of the primary photo-current. some semi-conductors, amongst which selenium occupies an important place, a large secondary current flows in addition to the primary photo-current. The origin of this secondary current is as yet obscure. It makes the material more conducting and arises from a change in the state of the material occasioned by the accumulation of primary electrons at the boundary. A definite time is required before the secondary current is fully established, leading to considerable time lag in the attainment of the final state for a given intensity of illumination. Selenium and thallous sulphide are particularly good materials for the construction of conductivity cells. The change in resistance which takes place on irradiation can be measured in the usual manner by a bridge circuit.

These cells require the application of a potential of about 100 volts. They are sensitive to temperature changes (one degree rise in temperature leads to a 2 per cent. change in response) and have a time lag of approximately one-tenth of a second between receiving the light and registering full change in resistance. The light sensitivity is not linear. The cells are very selective with a sharply marked maximum in the red or infra-red. This makes them very much more red sensitive than the human eye. They cannot be used at all for accurate work or for measurement but are best used merely for the

detection of light.

Light counters

When extremely weak sources of light are to be detected photo-counters are used. Ordinary photo-cells are more sensitive than the eye and in the red the conductivity cells are very much more sensitive than visual detection. For white light the visual threshold is 10-12 lumen. A sensitive cell gives a current of 6×10^{-6} amp, with this intensity of illumination if it has the same aperture as the eve (a disc of 6 mm. diameter). Such a current can be detected with the aid of amplifiers. A typical cell has a much greater receiving area than the eye and is correspondingly more sensitive if a wide aperture beam is available.

A light counter, or photo-counter, is a modified form of a Geiger-Müller counter, the application of which to radioactive observations will be discussed later. In the Geiger-Müller counter a metal cylinder (cathode) is enclosed in a glass tube. Down the cylinder axis passes a clean wire (anode). The tube is filled with gas at about a pressure of 5 cm. and a potential applied just below the spark breakdown potential. The production of ionisation near the wire by any means leads to ionisation by collision with resulting breakdown and the passage of a large current which can be detected readily.

In a light counter the inside of the cylinder is coated with a photo-sensitive material and the axial wire with a highresistance film. The applied potential exceeds the breakdown potential for a clean wire but no discharge takes place because of the high-resistance film. When light falls upon the sensitive cathode photo-electrons are released. They ionise by collision. Owing to the high-resistance coating on the wire a single pulse of current is produced for each photo-electron liberated.

The photo-counter is essentially a device for the detection of extremely weak light sources. With a detecting area of 12 sq. cms. it is possible to detect a light flux of only a single quantum per second. The remarkable nature of this will be appreciated when it is recognised that a quantum of yellow light (sodium light) has an energy content of only

 3.3×10^{-12} erg.

REFERENCES

[&]quot;Photo Electricity." H. S. Allen. (1925.)
"Photo Cells." Campbell and Ritchie. (1934.)
"Electron Emission and Absorption Phenomena." J. H. de Boer. (1935.)

CHAPTER 8

ATOMIC SPECTRA AND THE PERIODIC TABLE

The spectrum of ionised helium

Helium (atomic number 2) has two electrons in its outer structure and a nuclear mass of about 4. If it is ionised by the complete removal of one of the electrons the residual atom possesses only one outer electron and is therefore like hydrogen, except that the nucleus has twice the charge of the hydrogen nucleus and about four times the mass. In Bohr's simple theory for the Balmer series it is assumed that the nucleus is so much heavier than the electron that to a first approximation its mass may be treated as infinite. An infinite nucleus would remain at rest, but actually the nucleus and the electron both rotate about their common centre of gravity. It can be shown from a simple treatment of the dynamics of rotating bodies that the formulæ derived on the basis of an infinitely massive nucleus still hold for a non-infinitely massive nucleus if the electron mass m in the formulæ is, throughout, replaced by the "reduced mass" μ where $\mu = mM/(m+M)$, M being the nuclear mass. When account is taken of the finite nuclear mass, interesting differences appear between the spectra of hydrogen and ionised helium.

The spectrum of ionised helium should be similar to that of hydrogen since both atoms have a single electron, but there

will be differences due to:

(1) The nuclear charge being 2 instead of 1;

(2) The nuclear mass being approximately four times that of hydrogen.

Considering first the effect of the increased charge, and neglecting the difference in nuclear mass by assuming both to be infinite, it is clear that ionised helium will give series similar to the Balmer series, the lines obeying the formula

$$\begin{array}{l} \nu' = & \{2\pi^2\mathbf{Z}^2e^4m/ch^3\}\{1/n_1^2-1/n_2^2\} \\ = & \mathbf{Z}^2\mathbf{R}_{\mathrm{H}}\{1/n_1^2-1/n_2^2\} \end{array}$$

 $R_{\rm H}$ is Rydberg's constant for the hydrogen atom. Since $Z{=}2$ the series of He⁺ will be identical in form with that of hydrogen but with $4R_{\rm H}$ replacing $R_{\rm H}$. To a first approximation this is experimentally confirmed since the alternate lines of the ionised helium series, discovered by Pickering, almost coincide

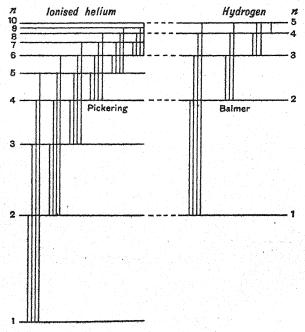


Fig. 8.1

with hydrogen lines as predicted by the Bohr theory (see Fig. 8.1).

The coincidence between the lines from the two atoms is, however, not absolutely exact, the cause of the difference being the different nuclear masses. The Rydberg constants for the two atoms are not quite identical when the reduced masses of the electrons are employed. The general expression for the Rydberg constant using the reduced mass (defined as for Z=1) is

$$\mathbf{R} = \{2\pi^2 e^4 m/ch^3\} \{1/(1+m/\mathbf{M})\}$$

If R^+_{He} and R_H are the Rydberg constants for ionised helium and hydrogen respectively and M_{He} and M_H the nuclear masses, we have from the above

$$\begin{array}{l} {\rm R^+_{He}/R_H}\!=\!(1\!+\!m/{\rm M_H})/(1\!+\!m/{\rm M_{He}}) \\ =\!(1\!+\!m/{\rm M_H})(1\!-\!m/{\rm M_{He}}) \end{array}$$

to a first approximation, since the ratio of the mass of the electron to that of a nucleus is small

Thus

$$R^{+}_{He}/R_{H} = 1 + (m/M_{H}) - (m/M_{He})$$

= $1 + (m/M_{H})(1 - M_{H}/M_{He})$

The quantity M_H/M_{Ho} is the ratio of the mass of the proton to that of the helium nucleus and is known from mass spectrographic date to be 1/3.98. Therefore an experimental determination of the ratio R^+_{Ho}/R_H leads to a value for m/M_H which is the ratio of the mass of the electron to that of the proton. The numerical experimental value for R^+_{Ho} is 109,722 and that for R_H =109,677. When these values are inserted into the above expression the ratio m/M_H is found to be 1/1843 which is very close to the value found from deflection experiments upon electrons. This was one of the early triumphs of the Rohr theory. So accurately is the formula obeyed that the spectroscopic method is now considered one of the best means for precisely measuring the electron mass and the ratio of the mass to the charge.

Series in line spectra

The series which appear in the spectra of hydrogen and ionised helium are obvious to the eye. Most spectra are considerably more complex than these two, and although many contain thousands of lines it has been found possible to arrange all the lines of any spectrum into series. In the large majority of cases the individual members of a series overlap into other series and a great deal of skill is required to separate the series members from the confused mass of lines usually observable. The spectra emitted by the many-valent atoms to the right of the Periodic Table, and by the rare earths, are particularly complicated. We shall therefore first consider a type of spectrum called "hydrogen like." Such a spectrum is one

emitted by a monovalent atom such as an alkali metal, or an alkali earth metal which has had one electron removed by ionisation. The spectra of all the atoms which occur in a given column of the Periodic Table are closely similar to each other in structure so that a discussion of the characteristics of the spectrum of sodium will cover the general details to be found in all hydrogen-like spectra.

Although the Bohr theory fits the grosser details of the hydrogen spectrum extremely well, refined observation shows that each member of the Balmer series is not a simple line but has a very narrow structure. Each line is complex, exhibiting what is called a "fine structure." The explanation of the origin of this fine structure is related to the explanation of the complex series found in all hydrogen-like spectra. Sommerfeld succeeded in accounting for both phenomena by postulating the existence of elliptical orbits in addition to the circular orbits proposed by Bohr.

Elliptic orbits

An electron in an elliptic orbit has two degrees of freedom. When represented in polar co-ordinates these are r, the distance of the electron from the nucleus, and the azimuthal angle θ . Each degree of freedom must be quantised employing a separate quantum condition. Bohr's quantisation of the angular momentum in a circular orbit was shown by Wilson and Sommerfeld to be a special case of a more generalised expression for quantisation which is

$$\int pdq = nh.$$

In this p is a momentum and q the corresponding position coordinate. Thus for a Bohr circular orbit the integral carried over a complete revolution gives

$$\int_{0}^{2\pi r} mv \cdot dx = nh$$
$$mvr = nh/2\pi.$$

i.e.

This is Bohr's condition of quantisation.

For an elliptical orbit, each variable obeys the Wilson-Sommerfeld relationship, therefore

$$\int p_{\theta} d\theta = kh$$

 $\int p_{r} dr = n_{r} h$

 p_{θ} and p_r are the angular and radial momenta. Two new quantum numbers, k and n_r , have appeared replacing the previous one, n. By integrating over a complete revolution it can be proved that

$$1 - \epsilon^2 = k^2/(k + n_r)^2 = b^2/a^2$$

where ϵ is the eccentricity of an ellipse with semi major and minor axes a and b respectively. Since both the quantum numbers k and n_r are integral, the sum $n=k+n_r$ can have values $1, 2, 3, \ldots$ The quantity n is called the total quantum number of the electron. It follows that only a certain number of possible orbits exist, namely, those in which the ratio of the major to the minor axis is the ratio of two whole numbers.

The evaluation of the total energy of an electron in an elliptical orbit by the method of Bohr shows that

$$\mathbf{E} = -2\pi^2 \mathbf{Z}^2 e^4 m/n^2 h^2$$

which is identical in value with that for a circular Bohr orbit with radius equal to a, the semi major axis of the ellipse. As far as this approximation is concerned, therefore, circular and elliptical orbits with the same n value have the same energy and therefore will not be distinguishable. Actually this is not exactly the case.

From the expression for the eccentricity of the ellipses given above, it will be seen that when n=k the path is circular since ϵ is equal to zero. The quantum number k can itself never equal zero since the ellipse then degenerates into a straight line passing through the nucleus. Clearly for a given value of n=k+n, k can take on n possible different values. There are therefore n types of ellipse, the orbits becoming more eccentric the lower the value of k. (According to this picture the angular momentum of the electron in its orbit is $kh/2\pi$, but detailed analysis has shown that the picture is not entirely correct: it is preferable to use a quantum number l=k-1, then the wave mechanics indicates that the orbital angular momentum is $\sqrt{l(l+1)} \cdot h/2\pi$.) Fig. 8.2 illustrates the type of elliptical orbit which the theory allows.

Although simple theory predicts that all orbits with the same major axis will have the same energy, Sommerfeld proved that

this is not strictly true. An electron in an elliptic orbit approaches close to the nucleus and acquires a high velocity in doing so, the velocity near the nucleus being greater than when farther off. The Theory of Relativity shows that when a particle increases its velocity its mass also increases. The effect is only appreciable at very high speeds but is sufficient to make the energy of the electron in a more elliptic orbit greater than that in a less elliptic orbit, even if the major axis remains unchanged. Thus of all the orbits with a given n value that

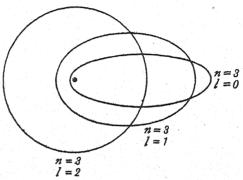


Fig. 8.2

with l=0 has associated with it a somewhat greater energy than that with l=1, that with l=1, a greater energy than that with l=2, etc. This will now be applied to the explanation of the details of hydrogen-like spectra.

One electron spectra

As will be discussed in more detail later, the structure of the Periodic Table enables us to conclude that the extra-nuclear electrons in atoms can be arranged in closed shells. Effectively the alkali metals consist of a nucleus, a number of completely filled closed shells and a single outer electron, the optical electron. The members of the next group of the Periodic Table have two optical electrons, those in the succeeding group three, etc. Sodium can thus be treated as an atom in which only one electron is involved in the production of the spectrum.

This optical electron can occupy different orbits and when in a particular orbit it is described by a *small* letter which identifies the orbit. When the electron is in an orbit with l=0 it is called an s electron, when in an orbit with l=1 it is called a p electron, etc. The particular letters employed arise from earlier empirical classification (see below). In Table I the small letters designate the electron. When the electron is in a certain orbit the atom is then said to be in a corresponding state, indicated by a capital letter. Here the same letter is used for the electron designation and the atomic state, but this only occurs in one electron spectra. In calcium, for instance, where there are two valence electrons, one of these may be a p electron, the other a d electron, and the atom may be in an F state.

TABLE I

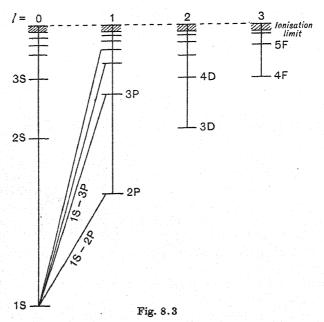
Orbital angular quantum number of electron	0 1	2 3	4	5 6
Electron designation	8 p	d f	g	h i
Atomic state	S P	D F	G	H I

Orbital designations are also shown in Fig. 8.3. The total quantum number n is placed before the letter indicating the l value of the orbit. An electron in the second orbit with l=1 is described as a 2p electron, one in the fifth orbit with l=4 is a 5g electron, etc. From what has previously been discussed it follows that the lowest possible n value associated with a given value of l is always l+1. Thus ls, 2p, 3d . . . are the lowest possible orbits with l values 0, 1, 2, . . . etc. The orbits lp, 2d, 3f, . . . etc., do not exist. On plotting the energies associated with the orbits a diagram like that in Fig. 8.3 is obtained. Each energy level, which corresponds to an orbit, is called a spectral "term." When the electron is in a particular orbit the atom is said to be in an atomic state which has the same notation as the term.

Transitions can take place between specified terms, and when the electron moves from one term to another, the atom radiates a *single quantum*, which is *one line* of the spectrum. The complex spectrum usually observed represents line transitions from a *large number of atoms*, since one transition means

one emission line. For any given value of l the terms run to a limit, the "series limit," such a group of terms being called a sequence.

The well known sodium yellow line was formerly described as 1S-2P, since it was treated as a transition from an upper 2P level to a lower 1S level. Both theory and observation show that the number of possible transitions is limited by a "selection principle," that is to say, only selected transitions can take place. Transitions are only possible between two



terms whose l values differ by ± 1 . This is usually written as $\Delta l = \pm 1$. In accordance with this rule transitions can take place between any S and any P term, for example, or any P and any D term, but not between an S and an S term (for then $\Delta l = 0$) or an S and a D term ($\Delta l = 2$), etc. It is clear that series similar to the Balmer series can arise by the transitions from all the terms in a sequence to one term in another sequence. Clearly, for example, 1S-mP represents a series, if m has the values 2, 3, 4, . . . etc.

Historically certain series were found and named before the

advent of the quantum theory of spectra. In sodium these were:

Sharp series .	•				. 2	P-mS
Principal series	•				. 1	S-mP
Diffuse series.			•		2	P-mD
Fundamental ser	ies				3	D-mF

The names were adopted mainly because of the appearance of the lines. The letters s, p, d, f, are the first letters of the names of the series. The above notation is a short form of writing down the term energies, since we know that the frequency of a line is given by the difference of two terms as in the Balmer series. Observation, in fact, shows that all series are built up in the same manner out of terms of the form R/n_2 in which R is Rydberg's constant, but, in the general case, n is not necessarily integral. Each term can also be written as $R/(n-\mu)^2$ where n is integral and μ is less than This quantity, μ , the quantum defect, represents the deviation from a whole number and is found to be practically constant for all terms in a sequence. The value of μ differs from sequence to sequence, diminishing with diminishing ellipticity. It is very small in f orbits (which have l=3), hence the latter give terms closely similar to hydrogen terms. For this reason these were called "fundamental," but we now know that there is no special significance to be attached to them. The total quantum numbers in these transition descriptions are not the same as those now generally employed. The above were purely empirical and are now usually replaced by the true quantum numbers shown in Fig. 8.4. The empirical description has however persisted in some publications, the quantum number adopted being based upon the "effective" number in the Rydberg sequence.

The spinning electron

Although the recognition of the existence of elliptical orbits with different l values succeeded in accounting, in a general manner, for the term combinations observed, the *multiplet structure* of spectra (sometimes called the fine structure) finds no explanation without the introduction of further electronic properties. Thus it has been known for many years that the yellow sodium line, and the remaining members of the whole

of its series, are close doublets. Close triplets are found for series members in magnesium, and as one moves to the right of the Periodic Table the members of series become more and more complex. The heavier the atom the greater is the wavelength separation of these multiplets. For instance the doublet components of the sodium yellow line are only 6 angströms apart, whilst the visible mercury triplet extends over 1,400 angströms. The name "fine structure" is clearly not applicable to the heavier atoms so that "multiplet structure" is to be preferred.

Since the quantum numbers n and l fail to account for multiplet structure, Goudsmit and Ühlenbeck in 1926 put forward the theory of electron spin. According to this every electron has a spin of $(1/2)(h/2\pi)$. Clearly a spinning electron must have a magnetic moment and as the orbital rotation of an electron also produces a magnetic moment these two will

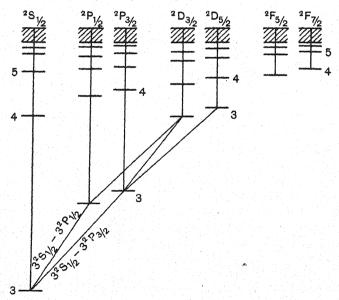


Fig. 8.4

interact. The spin angular momentum, denoted by the quantum number s ($s=\frac{1}{2}$) can combine vectorially with the orbital angular momentum l to form a resultant denoted by the quantum number j, which thus represents the total angular

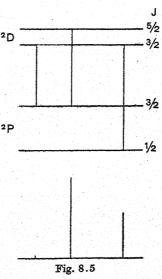
momentum of the electron. The vector addition is usually shown as $j=l\pm s$, the addition being quantised.

Consider now the effect of electron spin upon the S, P, D, and F terms of sodium. The S terms each have an l value of zero and as $s=\frac{1}{2}$, j must be equal to $\frac{1}{2}$. The j value is always indicated by a suffix, the term now being described as $S_{\frac{1}{2}}$. The P terms each have l=1, hence the vector addition of s and l produces $j=\frac{1}{2}$ when the vectors oppose each other, and j=3/2 when they act parallel. What was previously treated as a single P term is now seen to split up into two levels, since the energies of the electron orbits differ in the two cases. The two terms are described as $P_{1/2}$ and $P_{3/2}$ respectively. In like manner $D_{3/2}$, $D_{5/2}$ and $F_{5/2}$, $F_{7/2}$ terms arise from the interaction between electron spin and orbital angular momenta for D and F levels. Apart from the S term, which alone remains single because l=0, every term is double. The terms are therefore

called doublets. This is indicated by a superscript 2 written to the left of the term, e.g. ²S_{1/2}, ²P_{1/2},

 $^{2}P_{3/2}$, $^{2}D_{3/2}$, $^{2}D_{5/2}$, etc.

When account is taken of the electron spin the transition 3S-3P (the yellow line) becomes the doublet shown in Fig. 8.4. A selection principle also operates in the multiplet term transitions such that $\Delta j = \pm 1$ or 0 (but the transition $0 \to 0$ is forbidden). The operation of the selection rule makes, for example, 3P-3D have three constituent lines, since the transition $\Delta j = \frac{1}{2} \to \frac{5}{2}$ is not permitted. The three lines are still said to form a doublet



multiplet since the terms involved have a doublet structure (see Fig. 8.5).

Many electron spectra

When an atom possesses a number of valency electrons the multiplicity becomes more complex. Each of these electrons has its own particular l value and also a spin of $\frac{1}{2}$. In different cases the spin and orbital momenta can combine differently. It is possible for the l and s of each electron to combine to give a j value for each. The j values then combine to form a total resultant J which represents the total angular momentum of all the valency electrons. (In the vector notation capital letters are used to represent resultant vector sums for the whole atom, small letters referring to individual electrons.) This type of vector coupling is called jj coupling. In an alternative arrangement the l values of all the electrons unite to form a resultant L and the s values unite to a resultant S. Then, as before, we have J=L+S. When there are a number of valency electrons, S can have more than one value for a given L since the individual s values can combine in various ways. This leads to more than one multiplicity. The manner in which this arises will be made clear by a consideration of a two-electron spectrum such as calcium.

With two valency electrons S can have the value 0 or 1 according to whether the two electron spins oppose or assist each other. In the first case, when S=0, J is, in every term, equal to L, the result being that all the terms are single valued, e.g. ${}^{1}S_{0}$, ${}^{1}P_{1}$, ${}^{1}D_{2}$, ${}^{1}F_{3}$, etc. The multiplicity (1) is written to the left, as in the case of the doublets, and it will be noticed that the J value is in each case the same as the L value.

When S=1 each term (other than the S terms which are always single) becomes triple. For example, a P term has, L=1 and as S=1 these two vectors can combine to give the quantum numbers 0, 1, 2 for J. The possible terms are then 3P_0 , 3P_1 , 3P_2 . In like manner a D term with L=2 gives resultant J values 1, 2, 3 when combining with S=1. The triplet D terms are thus 3D_1 , 3D_2 , 3D_3 . All the remaining terms of the spectrum split up into triplets. It is easy to show that the multiplicity is in every case 2S+1. When S=0 we have singlets, when $S=\frac{1}{2}$ we have doublets, triplets when S=1, quartets when S=3/2, etc. An atom possessing, for example, seven d electrons in the outer shell can have S values up to

7/2 leading to octet multiplicity. More complex multiplicities also exist.

The interval and intensity rules

The levels constituting a multiplet in many cases are found to obey fairly accurately a rule called the Landé interval rule. This states that the interval in frequency between two terms with total angular momenta J+1 and J respectively is proportional to J+1. Take the case of a 3D multiplet, the terms in which are 3D_3 , 3D_2 , 3D_1 . In this, the levels are spaced so that the frequency-intervals between them are in the ratio 3:2. A 4D multiplet with $^4D_{7/2}$, $^4D_{5/2}$, $^4D_{3/2}$, $^4D_{1/2}$ terms would exhibit interval ratios of 7:5:3. There are many deviations from the interval rule but these can be accounted for. They usually occur when the electron coupling is neither completely of the jj type or the LS type, but is intermediate between these two extremes.

The line transitions forming a multiplet are generally of different intensities. It can be shown both theoretically and experimentally that the sum of the intensities of all the lines coming to, or starting from, any given $\bf J$ term of a multiplet is proportional to $2\bf J+1$. Applying this, for example, to the two members of the yellow sodium doublet $3^2S_{1/2}-3^2P_{1/2}$ and $3^2S_{1/2}-3^2P_{3/2}$, shows that the intensity ratio of the two lines should be 1:2. This is indeed the case as confirmed by observation.

The Pauli exclusion principle.

One of the most outstanding successes of the spectroscopic approach to the study of atomic structure is the manner in which the Periodic Table of the elements can be accounted for. It has already been pointed out that an electron requires the quantum numbers n, l, s, and j to define its orbit. There are other ways of looking at this. If the emitting atoms are maintained in a moderately strong magnetic field each single spectrum line splits up into a small group of lines usually symmetrically spaced. This is the Zeeman effect which has been known for many years. Each term is found to split up into 2j+1 magnetic levels. Physically this means that in a

magnetic field the vector j can only take up distinct quantised directions with reference to the external magnetic field which is being applied. It must be concluded that a further quantum number is required. This is called the *magnetic quantum number m* and clearly exists theoretically even if there is no

magnetic field.

Suppose we consider a single electron only with its l and s values combining to form a resultant j. The external magnetic field can be increased until it is so strong that the magnetic coupling between l and s is broken, since the latter are coupled magnetically. There is then no resultant j and we now have instead of m the magnetic quantum numbers m_l and m_s which are respectively the projections of l and s in the direction of the magnetic field. When this takes place a Paschen-Back effect is said to have set in (after the discoverers). An electron can therefore more precisely be defined by the quantum numbers n, l, s, m_l , m_s , for the magnetic quantum numbers are possible properties even when there is no field. Since for every electron the spin, $s=\frac{1}{2}$, is the same, it follows that any electron can be distinguished by the four quantum numbers n, l, m_l , m_s that it would possess if it were in a strong magnetic field.

The possible four quantum numbers which an electron can acquire are governed by a fundamentally important rule, the Pauli Exclusion Principle. This states that no two electrons in an atom can have the same four quantum numbers n, l, m_l , m_s . When this principle is applied numerically to the electrons in atoms two important deductions can be made. Firstly, as a result of this limitation, it can be shown that in an atomic electron configuration there can be no more than $2n^2$ electrons with the same quantum number n. Secondly, it can be proved that for any given n value there can only be 2(2l+1) electrons with any particular value of l. These two deductions account completely for the form of the Periodic Table in the following manner.

The Periodic Table

Consider first all the possible electrons with n=1 that can exist in an atom. Since the maximum possible number is $2n^2$ there can only be two. It is clear also that l must be zero for both of these electrons. There are therefore two possible 1s

electrons, and no others with n=1. All the electrons which have the same total quantum number n are said to belong to a shell, the shells being lettered K, L, M, N, etc., for n=1, 2, 3, 4, etc. The completed first shell, the K shell, consists then of two 1s electrons.

For the second or L shell, the total quantum number is n=2 so that the shell can contain up to eight electrons $(2n^2=8)$. The second rule shows that this shell can contain no more than two 2s electrons and no more than six 2p electrons $\{2(2l+1)=6$ when $l=1\}$. As a result of the second rule sub-groups form within each shell, all the electrons within a sub-group having the same n and l values. These groups, first classified by Stoner, are shown in Table II for the K, L, M, N shells. All the electrons belonging to a sub-group are called "equivalent electrons," and it is usual to denote the number of such electrons in an atom by an index. Two 1s electrons are written as $1s^2$, four 2p electrons as $2p^4$, seven 4f electrons as $4f^7$, etc.

TABLE II

Shell	K	L	М	N
n	1	2	3	4
Possible electrons	ls	2s 2p	3s 3p 3d	4s 4p 4d 4f
Maximum of electrons .	2	2 6	2 6 10	2 6 10 14
Total 2n ²	2	8	18	32

The structure of the Periodic Table will now be considered in the light of the possible numbers of electrons in the subgroups. The fundamental assumption made is that the number of electrons in an atom equals the atomic number, i.e. the numerical position of the atom in the Periodic Table, starting with hydrogen as the first atom. The chemical properties of an atom are completely determined by the number of outer electrons which in turn is fixed by the nuclear charge. Any atomic electron structure is formed by the addition of a single electron to the electronic structure (or electronic configuration, as it is often called) characteristic of the preceding atom in the Periodic Table. The quantum numbers chosen for the

added electron are such as to place this electron into the most tightly bound state possible, reducing thereby the potential energy of the atom to a minimum. The atom is said to be in its normal state when the potential energy is a minimum.

Starting with hydrogen, which has one extra-nuclear electron, it is clear from Table II that this electron in the normal state occupies a 1s orbit. (It can occupy any of the higher orbits when the atom is excited.)

Helium has two electrons, both of which can go into 1s orbits. By doing this the potential energy is at a minimum. The configuration of normal helium is then 1s², thus the K shell is closed. A new period should therefore begin again with lithium, as indeed it does.

The third electron required for the formation of the lithium atom goes into a 2s orbit (to keep the potential energy again a minimum), so that the configuration for this atom is 1s² 2s.

Berylium, the next atom in the Periodic Table, has the configuration $1s^2$ $2s^2$. The K shell is therefore completed and the selectron sub-group of the L shell is also filled.

The next atom is boron, and as the 2s sub-group is now filled the added fifth electron must go into a 2p orbit so that the electron configuration becomes $1s^2$ $2s^2$ 2p. So the process continues until at neon the whole of the L shell is completed and a new period begins at sodium with a configuration which is $1s^2$ $2s^2$ $2p^6$ 3s, the 3s electron being the single valency or optical electron. In Table III the configurations of the first nineteen atoms are shown. This can be continued until all the Periodic Table is completely accounted for.

It will be noticed that at helium, neon, and argon the K, L, M shells are completed respectively (strictly, with argon only the first two sub-groups of the M shell are completed) and after each a new period begins, exactly as required by chemical data. The completion of a shell means chemical inertness, since valency is really determined by the degree of incompletion of the outer shell.

Not only are the *chemical* properties of the atoms completely accounted for, but this arrangement of sub-groups explains all the observed *spectra* of the elements. Take for example the alkali metals. The Table shows that each has closed shells and a single outer electron occupying the 2s, 3s, 4s, etc., orbits

for lithium, sodium, potassium, etc. All have the same valency, all should have similar chemical properties, and all should have similar doublet spectra, as is the case.

It is also clear that if an atom is ionised its spectrum should resemble that of the atom directly preceding it in the Table. This is invariably the case and has been verified for a large number of atoms.

Theory shows that the 4s orbit is more tightly bound than the 3d orbit. As a result of this, when potassium is reached

		K	L	М	N
Atom	Atomic number	18	2s 2p	3s 3p 3d	4s 4p 4d 4f
H He	1 2	1 2			
Li Be B C N O F	3 4 5 6 7 8 9	2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	1 2 2 1 2 2 2 3 2 4 2 5 2 6		
Na Mg Al Si P S Cl A	11 12 13 14 15 16 17	2 Neon	2 6	1 2 2 1 2 2 2 2 2 3 2 4 2 5 6	
К	19	2	2 6	2 6	1

the nineteenth electron goes into a 4s rather than into a 3d orbit. This accounts for anomalies in the spectra of the elements following argon.

Hyperfine structure and nuclear spin

When individual spectrum lines are examined with very high resolving power it is found that many of them exhibit a very close complex structure. This is now called "hyperfine structure." The name "fine structure" has also been used,

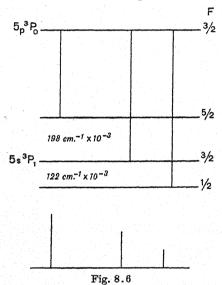
though the origin of the effect is quite distinct from that of the "fine structure" of the Balmer lines. It can be proved that the hyperfine structures are always due either to the existence of different isotopes in an element or else to the fact that the atomic nuclei themselves have a spin. The study of hyperfine structure has recently led to important information about the internal constitution of atomic nuclei. We shall consider briefly the effect of a nuclear spin upon spectrum lines.

Since nuclei are positively charged a spinning nucleus will have a magnetic moment. The spin of the nucleus must be quantised and can be represented by a quantum number I, the nuclear spin (the nuclear angular momentum) is therefore $Ih/2\pi$. Investigation proves that all the atoms with odd atomic weight have half integral values of I, ranging from $\frac{1}{2}$ to 9/2. It is fairly certain that atoms with mass numbers which are multiples of 4 have zero nuclear spin, but the remaining even atomic weight isotopes need not necessarily have a zero spin.

The magnetic moment of a spinning nucleus couples with the magnetic moment of the optical electrons which the latter have in virtue of their total angular momentum J. The interaction between the nuclear and electronic magnetic moments takes place in a quantised manner. The final result of this is that what was previously considered to be a single spectroscopic term splits up into a group of hyperfine structure levels. Owing to the relatively large nuclear mass, the nuclear magnetic moment is very much less than the electronic magnetic moment so that the splitting of the term due to nuclear magnetic moment is very small, hence the name hyperfine structure. It is apparent that what was formerly treated as a single line transition between two terms now becomes a small-scale multiplet since each term has a multiplet structure. The following particularly simple case will illustrate what takes place.

It can be shown that the nuclear spin I of arsenic is 3/2. Consider the effect of this upon the arsenic line $5s^3P_1-5p^3P_0$. For any term the spin couples with the total angular momentum J to form a set of hyperfine structure levels with quantum numbers F such that F=I+J. This is strictly analogous to the formation of an ordinary multiplet as represented by

J=L+S. The lower $5s^3P_1$ term has J=1 which can combine with I=3/2 to give three F values, namely, F=1/2, F=3/2, F=5/2. The positions of the levels in a hyperfine structure multiplet are such that the Landé interval rule is obeyed. The upper term $5p^3P_0$ has J equal to zero, hence F is single valued and equal to I. A selection rule operates, namely, $\Delta F=0$, ± 1 (with the transition $0\to 0$ excluded). Hence the complete hyperfine structure of the line under consideration, taking account of intensities which in this case are proportional to 2F+1 for each lower level, is as shown in Fig. 8.6.



An actual photograph of this hyperfine structure is shown in Plate IIB which is a reproduction of successive orders of the triplet hyperfine structure as given with a Fabry-Perot interferometer. The repeating order is marked in black. The whole pattern is about 0.1 angström wide so that it is quite clear that very sharp lines and very high resolving power are needed for the study of hyperfine structure.

Isotope effects.

The presence of a number of isotopes in an element can have different effects upon the hyperfine structures. The first influence of isotopes is the simple mass effect already discussed for the case of ionised helium. The slightly different masses of the isotopes lead to different Rydberg constants for each; as a result of this the series lines of each isotope are displaced slightly from each other. This effect has been observed only in the very light elements. For example, when the red neon lines are examined with high resolving power each is found to consist of two components with relative intensities 9:1. These correspond to the neon isotopes 20 and 22 whose abundance ratio is 9:1. In effect, the isotope hyperfine structure is able to show the existence of isotopes in the same way as

the mass spectrograph.

When an element (e.g. platinum) has both even and odd isotopes the former are unlikely to exhibit hyperfine structure due to nuclear spin since this is almost entirely a property of atoms with odd atomic weight. On the other hand, the odd isotope has a spin and therefore a hyperfine structure. As a result, a pattern due to the old isotope appears and at the optical centre of gravity of this we should expect a strong additional line representing the contribution of the even isotopes. However, in heavy atoms like platinum, mercury, etc., it is often found that the even isotope lines do not all fall exactly together at the optical centre of gravity of the pattern of the odd isotope, but are slightly displaced from each other. It is probable that this is due to different nuclear volumes for different isotopes. The study of isotope displacements leads to knowledge concerning the relative volumes of the nuclei of isotopes.

Recently it has been found that certain hyperfine structures fail to obey the interval rule and it has been proved that this arises through the particular nucleus involved not being spherically symmetrical. It is therefore evident that the study of hyperfine structure is of great value since from it information is derived about (1) nuclear spins, (2) nuclear magnetic moments, (3) nuclear volumes, (4) nuclear symmetry, (5) isotope abundance ratios.

REFERENCES

[&]quot;Introduction to Atomic Spectra." H. E. White. (1934.)
"Fine Structure in Line Spectra and Nuclear Spin." S. Tolansky. (1934.)

CHAPTER 9

X-RAYS AND THEIR PROPERTIES .

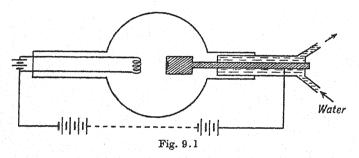
The production of X-rays

In 1895 Röntgen discovered X-rays, which are generated when a beam of fast cathode rays strikes a solid target. The major portion of the kinetic energy of the electrons is dissipated in the form of heat, the remaining small fraction producing X-rays which are electromagnetic waves of short wavelength. The nature of the emitted X-radiation depends upon the voltage applied across the cathode ray discharge tube, the amount of residual gas present, and the constitution of the solid target which is struck by the rays. The latter, called the anti-cathode, is usually a massive metal electrode suitably placed to receive a concentrated beam of fast electrons. The radiation from an anti-cathode is usually mixed, consisting of a continuous or "white radiation" upon which is superposed a line spectrum containing relatively few lines.

The earlier types of discharge tube were low-pressure gas discharge tubes, containing only a trace of gas, and upon applying a high potential across this, a stream of cathode rays resulted. This was directed on to an anti-cathode which then radiated. The voltage which can be maintained across such a tube depends in a critical manner upon the gas pressure, a factor which is very variable when the pressure is low. It was difficult with such a tube to maintain constant both the quality and the intensity of the X-radiation for any reasonable period of time. This created difficulties in measuring the effects due to the radiation. The introduction of the Coolidge tube overcame most of these troubles (see Fig. 9.1).

The distinction between this and the earlier tube is that here the vacuum is so extremely hard that even the application of a potential difference of 100,000 volts will not cause a current to pass. The source of the electrons is a heated

filament of special design. The thermions liberated from this filament are accelerated by means of an applied potential of the order of 50,000 volts and directed on to an anti-cathode consisting of a small piece of molybdenum embedded in a massive copper rod. (The molybdenum is replaced by another metal for some purposes.) The copper rod is often water cooled since there is a considerable development of heat. The amount of thermionic emission, which, of course, depends only upon the filament temperature, controls the intensity of the X-radiation produced, whilst the quality can be critically controlled by varying the applied voltage. Owing to the absence of gas, the potential can be maintained constant with little difficulty. Many modifications of the Coolidge tube have been designed. Some tubes are demountable to enable rapid inter-



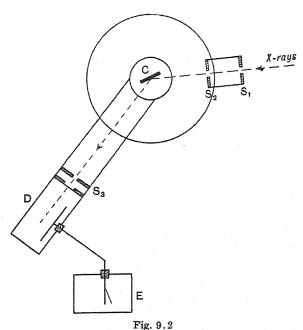
change of filament or anti-cathode. Since the anti-cathode becomes pitted after long usage, some tubes have a special device whereby the anti-cathode can be slowly rotated to reduce the effective time that any point is exposed to bombardment.

The form of the emission curve

Although X-rays are electromagnetic waves capable of producing interference and diffraction effects, the wavelengths encountered are so much shorter than that of ordinary light that an optical grating used in the normal manner is not suitable for studying X-rays. As will be shown in detail in the next chapter the atoms in crystals are arranged in a regular order and are sufficiently close to each other to enable one to use a crystal as a diffraction grating for X-rays. By the use

of this method, introduced and developed by Bragg, X-ray wavelengths can be accurately measured.

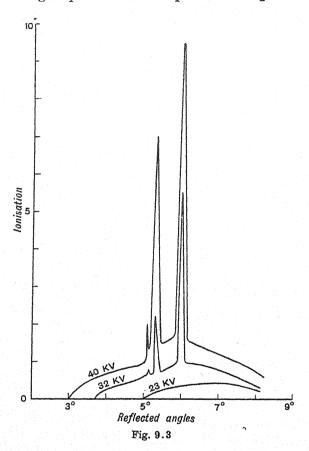
The Bragg X-ray spectrometer is illustrated in Fig. 9.2. The X-rays from a tube are collimated into a narrow beam by the slits S_1 and S_2 cut in lead plates, and then pass on to the crystal C which acts as a diffraction grating. The angular position of C is read by means of a vernier. After diffraction the rays enter an ionisation chamber D, filled with methyl iodide which strongly absorbs X-rays. The electrometer E



records the intensity of the ionisation in D. As will be shown later the wavelength can be derived from the angle of diffraction and a knowledge of the structure of the crystal. In addition to measuring wavelength, the ionisation spectrometer also measures the intensity for any particular wavelength.

When a plot is made of the intensity of X-ray emission against the wavelength, curves similar to that in Fig. 9.3 are obtained. This particular curve represents the emission from a target consisting of rhodium. The three curves correspond

to the emission observed when 23, 32, and 40 kilovolts respectively, are applied to a tube of the Coolidge type. In each case there is a continuous spectrum beginning at a definite wavelength which depends only upon the applied voltage. When 23 kilovolts are used there is a continuous emission, but with higher potentials a line spectrum emerges. The lines

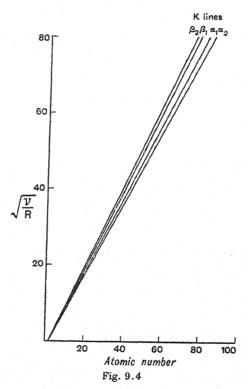


are characteristic of the metal constituting the anti-cathode.

The two main peaks in the intensity distribution are called the K-lines of the X-ray spectrum of the anti-cathode material. The K-radiation of every element consists of two lines (each of which is actually complex). In addition to this, each element emits softer radiation (i.e. longer wavelength radiation), the L-rays, consisting of three complex lines L_{α} , L_{β} , L_{γ} , and still softer emission, M, N, etc.

Moseley's law

In 1913 Moseley discovered an important law relating X-ray spectra with atomic number. He was led to enunciate this



law after examining the X-ray spectra of a large number of elements. He found a simple relationship between the frequency of the K lines of an element and its atomic number N. The latter is the numerical position the element occupies in the Periodic Table. On plotting the square root of the frequency of the K lines against atomic number, for the K_{α} and K_{β} lines, respectively, curves are obtained which are

nearly straight lines. The results obtained by Moseley are shown in Fig. 9.4. The K_a and K_{β} lines are each double.

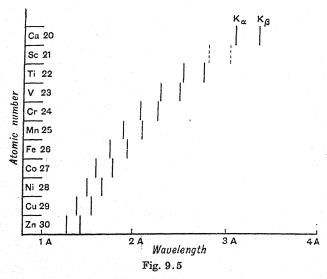
The wave numbers of the K_a lines are given almost exactly by the relationship

 $\nu = (3/4)R(N-1)^2$

where R is the Rydberg constant of optical spectroscopy. This formula can be rewritten

$$\nu = R(N-1)^2(1/1^2-1/2^2).$$

Clearly this expression is simply that for the Lyman series of a hydrogen-like atom with nuclear charge (N-1). In plotting



the members of the K series, jumps occurred at various places where elements were missing in the Periodic Table. Many of these missing elements have now been found and identified by means of their X-ray spectra. Fig. 9.5, given by Moseley, shows that nickel follows cobalt. This is in accordance with chemical properties, although the atomic weight of cobalt exceeds that of nickel. It was clear, in fact, that the atomic nuclear charge is more important than the atomic weight, a conclusion confirmed later by the discovery of isotopic constitution.

Moseley concluded from his investigations that the relationship between N and ν definitely proves the existence in the atom of a fundamental quantity which increases by regular steps up the Periodic Table. This quantity can only be the positive charge on the nucleus.

The X-ray spectra originate in the inner closed electronic shells of atoms and do not therefore exhibit the periodic effects characteristic of optical spectra, such as alternating multiplicities, etc. They are very much simpler than optical spectra, which, with many valent atoms, are highly complex. This is clear from Figs. 9.4 and 9.5. The striking simplicity and similarity of the spectra of the different atoms is quite apparent.

Energy levels and X-ray spectra

Since X-ray lines obey a law strictly analogous to that governing optical spectra, they must arise from electron transitions between two terms. The critical energy required to

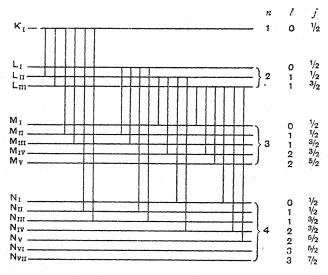


Fig. 9.6

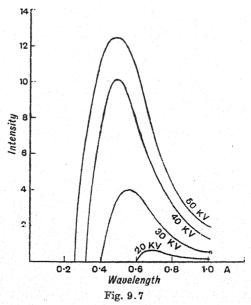
excite the K lines is just that needed to eject an electron from the innermost K shell. The vacant space left behind is filled by an electron falling into it from a higher orbit. As in the optical case, the electron transition results in the emission of a quantum of radiation, a K line. In like manner the L spectrum is excited when one of the electrons belonging to the

completed L shell is ejected, and so on. It is clear that a definite amount of energy (greater than hv) must be absorbed by an atom before the K-radiation of frequency ν is emitted.

Detailed investigation shows that the K level is single, the L level triple, the M level quintet, etc. The multiplicities of these levels explain fully the complex structures of the individual K_a , K_β , etc., lines. As an example of the levels found Fig. 9.6 illustrates the K, L, M, N energy levels for X-rays for uranium. The different levels are lettered K_I , L_I , L_{III} , etc., the n, l, j values for each being shown to the right. The usual selection rules of optical spectra apply, reducing the complexity of the individual compound lines. The vertical lines show the K and L line transitions.

The continuous X-ray spectrum

The above discussion has been concerned with the characteristic line spectrum of the anti-cathode. The properties of



the background continuous spectrum are quite different. As the potential across an X-ray tube is reduced, the K-radiation weakens, disappearing when the voltage is less than that

critical value giving the incident electrons an energy equal to the ionisation energy for the K shell. The continuum, however, is still emitted. From Fig. 9.3 it can be seen that the intensity drops to zero on the short wavelength side. This takes place at a limiting frequency given by $eV = h\nu_m$ where V is the potential applied across the tube. Although all the incident electrons have energy Ve there is a continuous emission degrading down to the long wavelength side instead of a single strong line of frequency ν_m because only very few of the electrons give up the whole of their energy at a single impact with one of the target atoms. Most of them require repeated atomic collisions before finally stopping, leading thus to a regular increase in wavelength. The effect of the alteration of the applied potential V upon the continuous emission alone is shown in Fig. 9.7. The total amount of this "white radiation" depends upon the nature of the anti-cathode, being approximately proportional to the atomic weight of the cathode material.

X-ray absorption

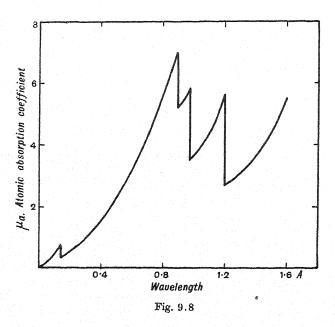
Like many optical spectra, X-ray spectra are often more easily studied in absorption than in emission. A beam of X-rays is weakened when passed through matter, a fraction being scattered away from the beam, and a fraction being absorbed in the photo-electric act of ejecting electrons from atoms. The scattering effect is only an appreciable fraction of the whole when the atomic weight of the absorbing material is small and the wavelength of the X-rays short. With monochromatic radiation, the intensity of a beam Io is reduced to I after passing through a thickness of material t so that $I = I_0 exp(-\mu t), \mu$ being called the linear absorption coefficient. As the absorption of X-rays depends upon the mass of the absorber this is usually rewritten as $I = I_0 exp(-\mu m/\rho)$ where μ/ρ is defined as the mass absorption coefficient, $m(=\rho t)$ being the mass per square centimetre of the absorbing screen. The coefficient μ/ρ can be divided into two parts $\mu/\rho = \sigma/\rho + \tau/\rho$, σ/ρ being the scattering coefficient (usually quite small) and τ/ρ

the transformation coefficient, so called since it is a measure of the amount of radiation absorbed by transformation into photo-electrons.

The absorption by a screen increases rapidly with the wavelength of the radiation and with the atomic number of the absorber. It is found that the basic variation is represented by an expression of the type

$$A \cdot \mu/\rho = \mu_a = CN^4\lambda^3$$

where A and N are the atomic weight and atomic number of the absorbing screen material, C being a constant. The quantity μ_a is the gram atomic absorption coefficient. However, the above law is not obeyed in a simple manner, as is evident from the absorption curve shown in Fig. 9.8.



The complexities of the actual case are explained by the fact that photo-electric absorption takes place in all electron shells of the absorbing material and because different shells are characterised by different values of the absorption energy.

Before an L electron can be ejected the quantum energy must exceed the appropriate absorption energy, that is the wavelength must be below a critical value. Similarly the wavelength must be lower still before a K electron can be ejected. The complete curve of μ_a against λ is therefore obtained by superposing curves, each with an approximate λ^3 form, but each cut off at a characteristic critical value of λ . At these values of λ μ suddenly falls (as λ increases) and the appearance of K, L, M . . . absorption edges follows.

The frequency of each absorption edge is a measure of the energy required to lift an electron from the corresponding X-ray level to the periphery of the atom, hence the difference in the frequency of two absorption edges gives the frequency difference between the two X-ray levels involved. By this means the term structure constituting an X-ray spectrum can be built up.

The Compton effect

X-rays can be scattered by matter in two quite distinct ways. In the first, coherent rays are scattered, without any wavelength change. This type of scattering leads to diffraction effects of which use is made in the crystal grating. The second type of scattering is incoherent, the scattered radiation having a longer wavelength than that of the incident rays. This is called the Compton effect, being first demonstrated and explained by Compton in 1922.

Compton directed a beam of X-rays on to graphite and measured the wavelength of both the incident radiation and of that scattered in different directions, using a Bragg ionisation spectrometer for the purpose. The scattered radiation was found to consist largely of the incident radiation but associated with this was some radiation with longer wavelength. The change in wavelength was found to depend upon the angle through which the radiation had been scattered.

This effect has been explained by Compton by attributing momentum to the incident quantum of radiation. When such a quantum strikes an electron the latter recoils, absorbing some of the energy of the radiation. The quantum is thus scattered with reduction in energy, leading to an increase in wavelength. Consider what takes place when a light quantum of energy $h\nu$ strikes a free electron. Such a quantum can be shown to have a momentum $h\nu/c$ where c is the velocity of light. As shown in Fig. 9.9 the incident quantum, after striking the electron, is scattered in the direction ϕ with reduced frequency ν' . The electron moves off with energy $m\nu^2/2$ in the direction θ . (Actually the electron mass is $m(1-v^2/c^2)^{-\frac{1}{2}}$ when the effect of increase of mass due to relativity is taken into account. Since

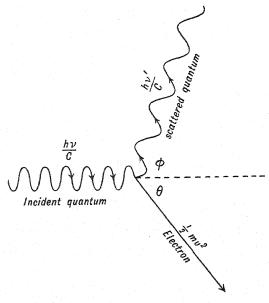


Fig. 9.9

the relativity correction cancels in the more exact analysis, it is neglected here.)

From the conservation of energy we have

$$h\nu = h\nu' + mv^2/2$$
.

Also, since momentum must be conserved, the forwards momentum equation is

$$h\nu/c = (h\nu'/c)\cos\phi + m\nu\cos\theta$$
,

and for the transverse components

$$(h\nu'/c)\sin\phi = mv\sin\theta$$
.

When these equations are solved, the change in frequency $\nu-\nu'$ which takes place when the radiation is scattered through an angle ϕ , can be evaluated. Converting into wavelengths, if $d\lambda$ is the wavelength change, it can be shown from the above equations that

$$d\lambda = (h/mc)(1-\cos\phi).$$

The numerical value of (h/mc) is 0.024 A, whence

$$d\lambda = 0.024(1-\cos\phi)$$
 angström.

This shows that the change in wavelength depends upon the angle of scatter ϕ , the maximum possible change being 0.048 A which occurs when $\phi=180^{\circ}$. The measurements made by Compton are in complete agreement with the above theory.

It is to be noted that the mass of the recoil particle is involved in the expression for the wavelength shift. The above value has been derived for a free electron. Strictly speaking there will be a Compton effect when any radiation is scattered by any particle. If the latter happens to be an electron bound in an atom its effective mass can be anything from about 2,000 to about 400,000 times that of a free electron. It is not possible to detect the resulting minute wavelength shift.

The refraction of X-rays

If X-rays are essentially similar to light waves they should suffer refraction on entering matter. Optical refraction is a consequence of the displacement of firmly bound electrons in the direction of a force applied by the incident electromagnetic wave, leading to refractive indices greater than unity. In X-rays the frequency of the wave is so much greater than that of light waves, that in all cases it exceeds the natural frequency of vibration of the electrons in matter. As a result, the electron displacements due to the radiation are opposite in direction to that of the force exerted, leading to a refractive index less than unity. This result is perhaps best seen with

the aid of Sellmeier's generalised formula for refractive index, μ , which is

$$\mu^2 = 1 + (e^2/\pi m) \sum_{1}^{N} n_s/(\nu_s^2 - \nu^2).$$

In this expression n_s is the number of electrons per unit volume, each with characteristic frequency ν_s , e and m are the electronic charge and mass, N the number of different kinds of electron (i.e. differently bound electrons), and ν the frequency of the incident light wave.

Clearly if $\nu > \nu_{\varepsilon}$ then μ must be less than unity. Furthermore, if $\nu > > \nu_{\varepsilon}$ we have approximately

$$\mu = 1 - ne^2/2\pi m v^2$$
.

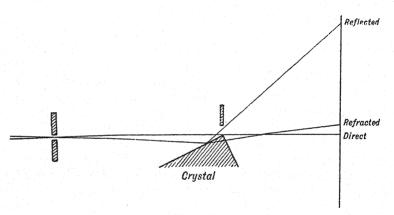


Fig. 9.10

Although this formula was originally derived classically, wave mechanics methods prove that it is still valid. Since ν is very high with X-rays, the refractive indices given by the above expression are only slightly less than unity. For example when the calculation is made, by substituting for a wavelength of 0.7 A, calcite is found to have a refractive index $\mu=1-1.84\times10^{-6}$. The very small difference between the refractive index of air and that of solids makes the measurement of the latter very difficult. Detection of deviation by a solid prism is not easy, but has been achieved by Siegbahn, Larsson, and Waller with the arrangement shown in Fig. 9.10.

The X-ray beam, falling at almost grazing incidence upon a quartz prism, suffers both refraction and reflection, as illustrated. The refractive index of the prism can be calculated from the positions of the images.

Reflection and interference of X-rays

Since for X-rays the refractive indices of metals are less than one, it is clear that at a suitable angle it is possible to produce total reflection, for the air is the dense medium and the metal the less dense medium. There will therefore be total external reflection, exactly analogous to the total internal reflection of optics. The critical glancing angle for total reflection can be shown to be $\theta = \sqrt{2(1-\mu)}$. Compton first observed total reflection in 1923. X-rays of wavelength 1.28 A were totally reflected from a glass mirror, the critical angle being small. Below this glancing angle 90 per cent. of the incident radiation is reflected.

The fact that so large a fraction of the incident beam suffers reflection at small angles, implies that within this limited angular range X-rays should exhibit interference phenomena similar to those observed optically. A ruled grating is in fact found to exhibit diffraction spectra, providing the angles of incidence and diffraction are close to grazing. Under such conditions, not only are the X-rays reflected, but also the number of effective lines per centimetre on the grating is considerably increased. Siegbahn has developed a special technique for the measurement of X-ray wavelengths with the aid of ruled gratings. Wavelengths can now be measured accurately both by crystal gratings and by mechanically ruled gratings, constituting thereby an important check upon the values derived for the spacings of atoms in crystals.

In addition to diffraction effects, direct interference can be demonstrated. Larsson, in 1929, obtained diffraction patterns with a slit 6μ wide, repeating well-known optical phenomena. Linnik has succeeded in obtaining direct interference fringes with a Lloyds mirror arrangement. Thin film interference fringes have been obtained by Kuhlenkampf in the following manner: A thin nickel film 1,400 A thick is deposited upon glass. X-rays are directed on to this at an angle slightly

exceeding the critical angle for nickel. A part of the radiation is reflected from the nickel-air interface and part penetrates and is reflected from the nickel-glass interface. These two beams unite, producing fringes which can be seen.

There is now a complete bridge between the properties of X-rays and of light, since the former show all the characteristic interference effects to be expected from electromagnetic waves with short wavelengths.

REFERENCES

[&]quot;The Crystalline State." Vol. 1. W. L. Bragg. (1933.)

[&]quot;Report on Progress in Physics." Vol. 1. (1934.) Physical Soc. of Lond.

CHAPTER 10

THE STRUCTURES OF CRYSTALS

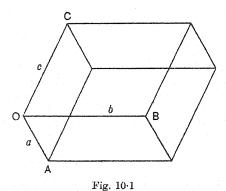
Introduction

The regular geometrical form of a crystal implies the occurrence of regular atomic or molecular arrangements in three dimensions. An atomic unit of pattern, the unit cell, is repeated and the crystal thus built up. A crystal will therefore possess an inherent periodic structure, the apparent form of which depends upon the direction of observation. In 1912 it occurred to von Laue that such an arrangement of atoms is essentially a three-dimensional grating and should therefore be able to produce diffraction effects with X-rays. This novel idea opened up an entirely new field of study and from its development both X-ray wavelength measurement and the study of crystal analysis evolved. At that time, existing ruled gratings were too coarse to be employed with X-rays.

Shortly after von Laue's discovery of the diffraction of X-rays by a crystal, W. L. Bragg showed how the structure of a crystal could be ascertained by using an elegant simplification of von Laue's method, and largely because of the work of this investigator a vast new field in the physics of solid bodies has been developed. The studies of crystal structure inaugurated by W. L. Bragg have both great scientific and great technical importance. An incidental development has led to high precision X-ray spectroscopy. It is clear that the conclusions to be drawn from an exhaustive study of matter in the solid state are of fundamental scientific importance. technical applications to engineering, chemistry, biology, textiles, etc., cover a very wide and rapidly growing field. Observers quickly found that the detailed study of complex crystals is difficult, a fact leading to highly specialised techniques. We shall review here only the general principles as applied to the simplest cases.

Crystal lattices and crystal planes

For many years crystallographers have classified crystals according to their external symmetry characteristics revealed by the angular relations between the crystal faces. Thirty-two classes were distinguished and named, each representing one of the possible geometric configurations of the centres, axes, and planes of symmetry which can exist. When the existence of these external symmetry requirements is combined with the assumption of an internal structure based upon a repetition of identical unit cells, one is led to a description of 230 possible space groups or ways of grouping symmetry elements. A

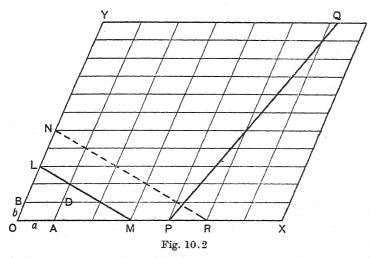


knowledge of the theory of space groups is essential to the analysis of the structures of complex crystals.

The three-dimensional unit cell upon which any crystal structure is based is defined by the lengths of its edges, a, b, c, and the angles between them. A typical cell is shown in Fig. 10.1, the three directions OA, OB, OC being called the axes of the crystal. The corners of the cell are lattice points, each point being occupied by an atom or molecule or group of molecules. It will be seen shortly that what is of importance in the X-ray study of crystals is the situation of crystal planes, that is, planes containing large concentrations of atoms. Any such plane is identified with respect to the crystal axes by a group of three numbers known as the Millerian indices (h, k, l) of the

plane in question. Such a plane makes intercepts on the axes a, b, c proportional to a/h, b/k, c/l. A two-dimensional model will make this clear.

Let the pattern in Fig. 10.2 represent a two-dimensional crystal built up by repetition of the unit cell of lattice points A, O, B, D, the lengths of the cell sides being a and b. Since we are dealing with a model in two dimensions instead of three we consider crystal rows instead of crystal planes. In the unit cell the corners only are occupied by atoms or molecular groups. Consider, for example, the row lying on the line LM. The intercepts of the line on the X and Y axes of the pattern,

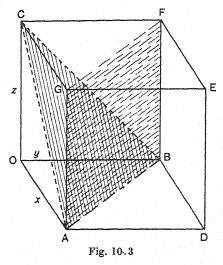


OM and OL, are simple integral multiples of the lengths a and b. The Millerian indices for this row are the *reciprocals* of these integral values reduced to their lowest terms. Thus in this case the lengths OM and OL are both the same number of integral multiples of the units a and b and the row is described as a \circ (11) row. The same designation clearly applies to any parallel row, e.g. NR.

Another example is represented by the row PQ. In this case the X intercept is one-fifth that of the Y intercept (in terms of the respective units a and b) and is negative, hence the row notation is $(\overline{\bf 5}\ {\bf 1})$. A minus sign is placed above the 5 as the row slopes in a different sense. By the same reasoning rows parallel to the Y axis are called $({\bf 1}\ {\bf 0})$ for then the Y

intercept is infinite, and rows parallel to the X axis are termed (0 1).

On extending the two-dimensional model to three dimensions it is clear that all crystal planes can be described by three Miller indices (h, k, l). An example is illustrated in Fig. 10.3 in which the unit cell is a simple cube. The nomenclature for all planes parallel to the diagonal plane ABC is $(1\ 1\ 1)$. That for planes parallel to the diagonal plane AGFB is $(1\ 1\ 0)$, since the intercepts are equal upon the x and y axes and infinite on the z axis. As further examples may be quoted the planes



parallel to OAGC which are (0 1 0), those parallel to ADEG which are (1 0 0), etc.

It may be noted that there are three possible forms of cubic lattice leading to different concentrations of atoms in special planes. These three types are the simple cube (above) with an atom (or molecule) at each corner, the body centred cube with an additional single atom at the cube centre, and the face-centred cube which is the simple cubic lattice plus an atom at the centre of each of the six faces.

X-ray diffraction by crystals

Laue was the first to realise that the close, ordered array of the atoms in a natural crystal would act as a diffraction grating

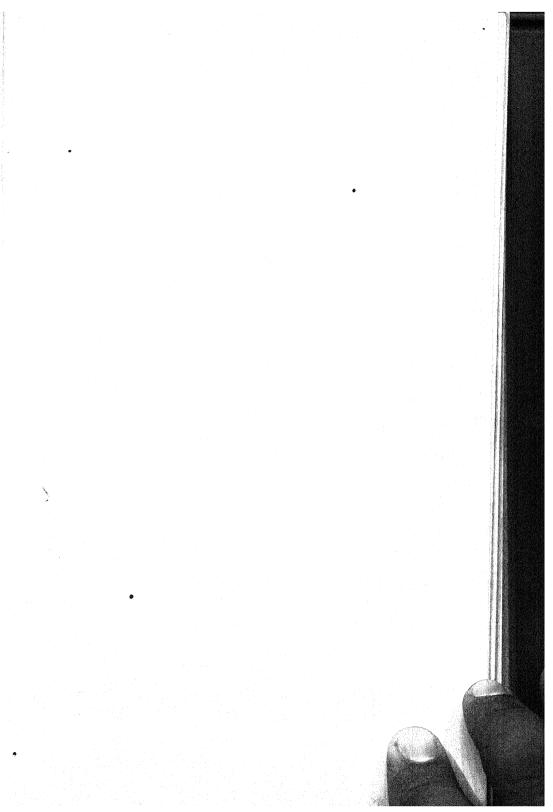
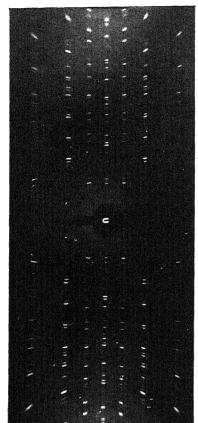
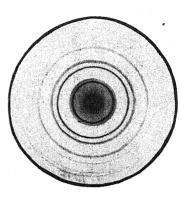


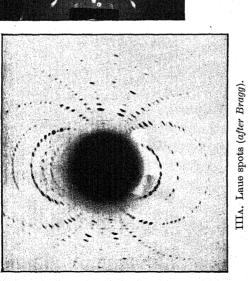
PLATE III



IIIB. Rotation photograph (after Bragg).



IIIb. Electron diffraction rings with thin film (after Bragg)



IIIc. Powder photograph, Debye-Scherrer type (after Bragg).



to X-rays if the latter were electromagnetic waves. In 1912 Laue, Friedrich, and Knipping passed a narrow pencil of X-rays through a crystal of zinc blende (ZnS) and then on to a photographic plate. The latter was found to exhibit an intense central image surrounded with a pattern of spots due to beams which had suffered deviation from the direct ray. Laue proved that the observed pattern could only have resulted from diffraction by a cubic arrangement of scattering centres. This experiment was of fundamental importance for it demonstrated conclusively that X-rays were short electromagnetic waves of wavelength some thousands of times smaller than light waves. A typical Laue photograph is shown in Plate IIIa. The Laue spots are only developed if the incident X-ray beam consists of "white radiation," i.e. is effectively a continuous band of wavelengths, as in the case of white light. The mathematical analysis of the positions of the spots is in general very complex.

A very elegant simplification in procedure was introduced by W. L. Bragg. In this treatment Bragg considers the X-rays to be effectively "reflected" from each crystal plane, according to the following mechanism. Suppose a beam of X-rays, this time monochromatic, not "white," passes through a crystal. Each atom scatters a very small fraction of the incident wave and in effect each becomes an emitter of secondary wavelets in a manner precisely analogous to that of the elements of the wave front in the Huygens construction for the propagation of light waves. The scattered wavelets will reinforce in any given direction only if they are in phase. Now all the scattered wavelets from the atoms in one plane will necessarily be in phase in that direction which bears to the direction of incidence of the X-rays the same relation as the direction of reflection bears to the direction of incidence when light falls on a plane surface. If, then, the radiation scattered from planes parallel to the plane under consideration also reinforces the contribution from that plane, in the direction in question, great intensity of scattering will result.

Consider parallel radiation, incident at glancing angle θ , reflected at the same angle from successive planes P_1 , P_2 , which are distant d apart. The retardation between the two beams is $2d \sin \theta$. If $2d \sin \theta = n\lambda$, where n is an integer, there

will be reinforcement for wavelength λ in the direction θ and in all other directions there will be destructive interference. This equation, known as Bragg's law, is the fundamental basis of crystal analysis. Assuming λ to be known, then measurement of θ gives d which is the spacing between the parallel crystal planes. By rotating the crystal to produce reflection from different sets of planes the shape and volume of the unit cell

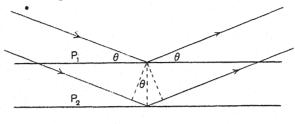


Fig. 10.4

can be derived. If the density of the crystal is also known, then from the volume of the unit cell it is possible to evaluate the *number* of atoms of various kinds in the cell.

Experimental methods of crystal analysis

(a) The Laue method.—The "white" radiation method of Laue is more difficult to apply and less powerful than methods using monochromatic radiation and the application of Bragg's law. It will therefore be briefly examined only. The observed spots have the following origin. With a fixed crystal and a fixed direction of incidence the various crystal planes constitute in effect a number of reflecting mirrors inclined at different angles to the incident beam. The emergent beam exhibits reinforcement only in the directions θ given by Bragg's equation. The angle is fixed for each plane, but when the incident radiation is white the appropriate wavelength for a particular angle is selected and reinforced. The result is the production of spots of reinforcement which exhibit clearly the basic symmetry of the crystal when this is suitably orientated with respect to the direction of incidence of the X-rays.

The interpretation of the Laue pattern is more difficult than the interpretation of the experimental data obtained by other methods which use monochromatic radiation (frequently the K doublet of molybdenum, two lines with wavelengths 0.7120 and 0.7076 A). We shall deal with these other methods now.

(b) The ionisation spectrometer method.—This method was developed by W. H. and W. L. Bragg for measuring reflections from crystal planes. The method is of historical interest since by it the first X-ray wavelengths were measured and the first crystal analyses made. The details were given in the chapter dealing with X-rays.

The crystal is successively turned through small regular increments, the detector being rotated through twice this angle. The particular plane involved in producing the reflection is deduced from the angular setting of the crystal and the direction of incidence.

Compared with photographic methods the ionisation spectrometer is laborious and slow, but it has the great advantage over these of giving accurate quantitative measurements of the intensities of the reflected beams without the intermediate use of microphotometers which always introduce errors. A knowledge of the relative intensities of reflections from different planes is necessary for complete analysis.

(c) The rotation-photograph method.—In this method, due to Schiebold and Polanyi, a small crystal is rocked or rotated slowly about a vertical axis by a clockwork motor. A narrow X-ray beam is directed horizontally on to the crystal, the diffracted beams being detected by a photographic film bent into a cylinder with the axis of rotation as the cylinder axis. The method is particularly suited to small crystals with linear dimensions of the order of 1 mm. The diffracted beams lead to small spots or lines upon the film, for as the crystal rotates successive planes are brought to the correct angle for reflection. The resulting photographs are similar to that shown in Plate Successive horizontal lines (layer lines) are produced when an important zone axis of the crystal is parallel to the axis of rotation. For example, if it is the c axis of the crystal that is parallel to the rotation axis, these represent reflections from the planes (h k 0), (h k 1), (h k 2), etc., for the respective successive layer lines. From the distances between the layer lines the space lattice can be derived.

(d) The powder-photograph method.—The powder method was devised independently by Debye and Scherrer and by Hull. In it the X-ray beam is allowed to fall upon a tube containing the crystal in a powdered form. The individual small crystals are oriented at random and amongst these are always some at the correct orientation for producing reflection from a given $(h \ k \ l)$ plane. Clearly such crystals may have the angle correct in all directions around the direction of incidence so that the beams diffracted by a given $(h \ k \ l)$ plane lie upon the surface of a cone. A ring is therefore produced upon a photographic plate which is set perpendicular to the incident beam. The concentric rings observed come respectively from different planes.

In practice it is found convenient to surround the tube containing the powder with a cylindrical film which thus records beams diffracted up to nearly 180°. On opening up such a film the picture obtained is similar to that shown in Plate IIIc. The curvature of the trace on the film changes as the angle of diffraction passes through 90°, hence circles of increasing radius of curvature appear as the diffraction approaches 90°, at which point straight lines are observed. The sign of the curvature then changes. For sharp lines it is necessary to have a fine powder, and a hair covered with fine crystals is often used. Since

then $\begin{array}{ccc} n\lambda = 2d \sin \theta \\ \Delta d. \sin \theta + d \cos \theta. \Delta \theta = 0, \\ i.e. & \Delta \theta / \Delta d = -\tan \theta / d. \end{array}$

As the angle of incidence θ approaches 90°, $\Delta\theta/\Delta d$ becomes very great so that small variations in d produce large variations in θ . This leads to high resolution, and slight differences in the lattice distance d can readily be detected for the rays reflected back nearly through 180°. Such small variations are of great significance in the study of metallic alloy crystals.

The powder method is widely applicable since any type of crystalline matter can be examined no matter how small the individual crystals may be. Furthermore, when crystal mixtures are studied each type of crystal records its own lines, and as these can usually be recognised it is possible to make estimates of the relative concentrations of the different

materials. The interpretation of the photographs is, however, difficult and so far only the simpler forms of crystal symmetry have been studied by this method.

It will be seen, therefore, that the different experimental methods which have been devised are best suited to different types of crystal. For a large single crystal the rotation method is most suitable since (apart from the fact that the estimated intensities are only approximate) a few photographs give all the information needed. The Bragg spectrometer method is laborious, but is precise and gives all the data accurately. It is used for large crystals when other methods fail to give the accuracy required. When the material to be studied is either finely divided or has a microcrystalline structure only the powder method can be employed. The Laue method has very limited applications.

The structures of KCl and NaCl crystals

In this section the structures which have been derived for sylvine (KCl) and rock salt (NaCl) will be discussed. These were the first crystal structures to be analysed by X-ray methods. Both crystals have cubic symmetry. Fig. 10.5 shows the ionisation spectrometer measurements made with the two crystals for the planes (100), (110), (111) using the radiation from a palladium tube. Two peaks appear in successive orders for each of the planes, due to K_{α} and K_{β} radiations.

Consider first the curves obtained with KCl. The peaks for the (100) face occur at glancing angles whose sines are in the ratio 1:2:3. They therefore represent the equation $n\lambda=2d\sin\theta$ for values of n=1, 2, 3, i.e. they are successive orders. Comparing now the angles at which the first order peak appears for the (100), (110), (111) planes respectively, we find the following. From Bragg's law $d \propto 1/\sin\theta$ so that comparison of the values of θ in the three cases gives the ratios of the spacings of the three planes. From the observations it is found that

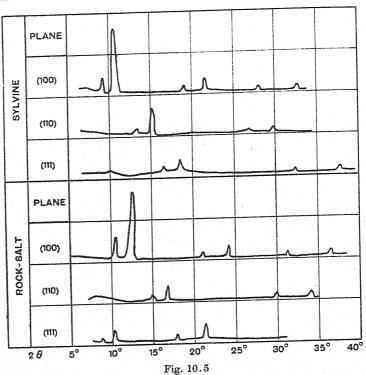
$$1/d_{(100)}: 1/d_{(110)}: 1/d_{(111)} = 1: \sqrt{2}: \sqrt{3}.$$

From Fig. 10.3 it is seen that these ratios arise from a simple cubic lattice structure of linear side d, since in such a unit cell

the (100) planes are d apart, the plane (110) is distant $\sqrt{2}.d/2$ from the point O, which contains the next (110) plane, and

the (111) spacing is $\sqrt{3}.d/3$.

Thus from the KCl measurements it was concluded that the crystal has a cubic structure, but no information was available as to the nature of the constituents at the lattice corners, i.e. are they molecules of KCl or ions of K and Cl alternately? The answer to this was given by the observations made on



NaCl. It is seen from Fig. 10.5 (111) that this crystal exhibits a weak reflection near to $2\theta = 10^{\circ}$ as well as that near to 20° . (In the curves the intensity is plotted against 2θ .) The latter is the peak corresponding to the first order of (111) in KCl. It is clear that the weak peaks near 10° are the true first order, and the value of the angle for this leads to $2d/\sqrt{3}$ for the $d_{(111)}$ spacing instead of the value $d/\sqrt{3}$ in the KCl crystal. Bragg

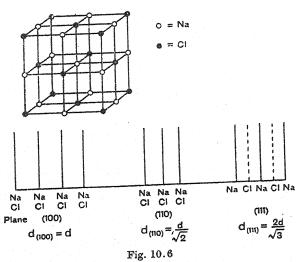
showed that the only way of explaining this was by assuming the lattice points to be occupied by ions and not by molecules. The intensity of scattered X-radiation is roughly proportional to the square of the atomic number of a scattering atom. and as the atomic numbers of potassium and chlorine are respectively 19 and 17 both scatter to approximately the same extent. The result of this is that the lattice appears to have identical diffracting centres at all the corners? In the case of NaCl the atomic numbers are 11 and 17, leading to a marked difference in the scattering from the different constituent atoms. On referring to Fig. 10.6 it can be seen that the vertical planes (100) contain equal numbers of Na and Cl ions, successive planes being identical. The same is true also for the (110) planes. But consider the case of the (111) planes and it will be seen that they consist of alternate parallel layers of Na ions and Cl ions. The spacing between a pair of parallel planes is $d/\sqrt{3}$ but what is measured is the true grating spacing. that between identical planes, i.e. $2d/\sqrt{3}$.

The relative weakness of the first order of the (111) spectrum comes about because the planes of the Na ions are exactly half-way between those of the Cl ions with the result that the rays reflected by the two are exactly out of phase and thus only the difference in amplitude between them is recorded. For the second order there will be reinforcement, as is clear by substituting in the Bragg formula.

The observed intensities are therefore completely accounted for by adopting the structure shown in Fig. 10.6, with alternate ions at the lattice corners instead of molecules. This structure can alternatively be viewed as built up out of interlocking face-centred cubic lattices of like atoms. Having determined the form of the structure W. H. Bragg and W. L. Bragg were able to decide upon the actual scale dimensions of the unit cell and also to evaluate the wavelength of the X-rays used.

The mass of each small cubic cell is $d^3\rho$ where d is the length of the side and ρ the density of the crystal. Each small cube effectively contains (for rock salt) half a sodium atom and half a chlorine atom, for there are eight atoms, one at each corner, but each of these is shared amongst eight cubic cells. Effectively the mass in each cell is half the molecular weight. Thus $d^3\rho = M/2$ where M is the weight of the NaCl molecule. This is $(23+35\cdot5)$ times the weight of the hydrogen atom, and as

the latter is known to be 1.65×10^{-24} gm., d can be found when the value for ρ (2·17) is inserted into the above expression. Substitution gives $d=2.814\times10^{-8}$ cm. For the first order reflection $\lambda=2d\sin\theta$ and on inserting the observed value of θ the wavelength found for the radiation is $\lambda=0.586\times10^{-8}$ cm., i.e. 0.586 A. This determination of wavelength for palladium radiation was the first wavelength measurement to be made in X-ray work.



By using the same radiation, the wavelength of which was now known, W. L. Bragg and co-workers undertook the determination of the spacings in more complex crystals. versely having evaluated accurately the spacing for rock salt, this crystal could now be used for the measurement of the wavelength of any other X-radiation.

More complex crystals

The NaCl structure may be considered as arising from the interpenetration of two face-centred cubic lattices, one of the lattices being displaced sideways (or upwards) by half its length. This is the simplest form of displacement possible. The case of zinc blende (ZnS) is a more complex variant of this. It is found that the Zn atoms all lie in one face-centred

cubic system and the S atoms in another. The lattices are, however, displaced in the manner exhibited in Fig. 10.7. The key to the determination of the position of the two sets of lattices is given by the relative intensities of the reflected spectra. Each S atom is at the centre of one of the small cubes. Each Zn atom is tetrahedrally surrounded by four S atoms and vice versa.

The diamond crystal has the same type of structure form as that of zinc blende but the atoms are identical in the two

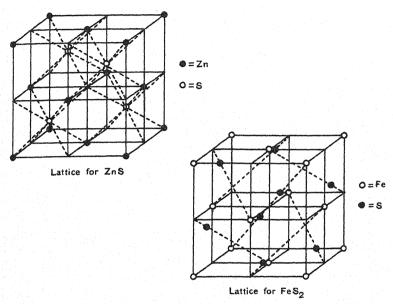


Fig. 10.7

displaced face-centred cubic systems. Each carbon atom of the one system is surrounded tetrahedrally by four carbon atoms of the other system.

In still more complex cases the interlocking systems are not symmetrically displaced. For example, in FeS₂ the S atoms are all situated one-quarter of the distance along the diagonal of a cubic element of Fe, not at the centre (see Fig. 10.7).

A knowledge of the intensities of the various orders from the selected planes is of the greatest value, being decisive in establishing the structures of the more complex crystals. It has therefore been found essential to examine in more detail the dependence of the intensity upon both the atomic number of a scatterer and upon the direction of scattering.

Factors affecting the intensity

Observation shows that the amplitude of X-radiation scattered by an atom is not directly proportional to the atomic number (i.e. the number of outer electrons) in a simple manner. The simple proportionality only holds at *small* scattering angles. The amount scattered falls off with increasing angle, the rate of fall, for a given wavelength, being greater with lighter atoms than with the heavier, since with the lighter atoms the average distance between the scattering electrons is greater.

According to classical electromagnetic theory, when a plane-polarised wave of amplitude A passes over and sets in vibration a free electron, the latter scatters a portion of the wave whose amplitude distant r from the electron in the plane of polarisation is (Ae^2/rmc^2) . An atom of atomic number Z contains Z outer electrons, and if each is assumed to scatter in accordance with the above expression, independent of its position, the resultant for the atom is Z times the above. As θ , the angle of scattering, increases the waves from different electrons become out of phase, the resulting interference reducing the resultant amplitude in the above case to $f \cdot Ae^2/rmc^2$ where f depends upon θ and is only equal to Z when the angle is small. This f falls off as θ increases. It is known as the atomic scattering factor.

When the scattering problem is treated from the viewpoint of wave mechanics analogous results are obtained, but with greater precision. Curves showing the relationship between f and θ can be constructed theoretically and are found to fit observations made upon the percentage of radiation scattered in different directions. Such f curves are used for calculating the intensities to be expected from different crystal planes.

Representation as a Fourier series

A strikingly elegant development in analytical technique has been introduced by W. H. Bragg. In this the relation between the periodic distribution of the scattering matter and the resulting X-ray diffraction can be expressed by regarding the crystal structure as represented by a three-dimensional

Fourier series. Every measurement of an X-ray reflection is treated as a measurement of the corresponding Fourier component in the crystal structure. The X-ray measurements give the intensities (but not the phases) of the diffracted beams. Instead of adjusting the positions of atoms by trial and error in order to account for observed intensities, the intensities of the reflections from different planes are observed and from these the electron density throughout the crystal is calculated.

The electron density distribution is calculated by Fourier analysis methods and when completed automatically shows up centres of atoms as concentrates of electrical density. The different types of atoms can be recognised by the degree of electrical concentration, which depends upon the atomic number. Such analyses are usually very difficult to carry out, but when results can be obtained they are very elegant, since a true contour picture of the electron distribution within the crystal is then obtained. The final analysis appears graphically in the form of a contour map, contour lines crowding in at high electron densities. In some analyses carried out with organic compounds molecular arrangements such as the hexagonal benzene ring appear clearly in the final diagram of the crystal structure.

It should be noticed that in treating complex crystals it is often found convenient to group together acid radicles (e.g. SO₄) and treat them effectively as units relative to the metallic components. In many organic crystals the whole organic molecule appears as the unit building stone of the crystal structure, since within the molecule the individual atoms are relatively much more tightly bound. This characteristic leads to a marked difference in the mode of approach to analysis of organic as compared with non-organic crystals. The positions of the molecules as structure elements can be found by the methods already described, but the location of individual atoms is extremely difficult. This is due partly to the complexity of the structures and partly to the fact that the amplitudes of scattering from the frequently met atoms of carbon, oxygen. and nitrogen are so similar that it is extremely difficult to distinguish one from the other. To this must be added the fact that hydrogen, which chemically is of such profound importance, by virtue of possessing only one electron, scatters

so little that, in general, its effects are not observed, at all. It will be clear that organic crystals present serious difficulties.

Interatomic forces

The results obtained from crystal analysis have greatly influenced the chemical theory of the solid state. The analysis shows whether in any particular case the aggregate consists of ions or complete molecules. Correspondingly the forces operating to hold together the constituent components of a molecular crystal are divided into two kinds, and we speak pictorially of heteropolar and homopolar bonds. We shall deal first with the heteropolar bond, sometimes called the ionic bond.

The crystal of NaCl represents a typical heteropolar case. In this, the electrostatic attraction between oppositely charged adjacent ions is the main binding force in the crystal. molecule is completely ionised even in the solid state, the act of solution leading, not to the production of ionisation, as was previously thought, but merely to the destruction of the lattice. The sodium atom has a single valent optical electron which is relatively loosely bound and the chlorine atom has an outer M shell which is incomplete, one electron being missing. Since an atom with an incomplete shell has a strong electron affinity the chlorine takes over the loose valency electron of the sodium with the resulting formation of the ions Na+ and Cl-, the radius of the Na+ atom being less than that of the Cl- atom. Although the ions attract each other strongly they are prevented from falling into one another by the operation of strong repulsive forces which are only considerable when the ions approach each other to within a certain close distance. (This conception is quite reasonable, for when the total electrical complex representing one atom begins to penetrate into that representing the other a force of repulsion may be expected. As an analogy consider two oppositely charged balls attracting each other. They will come into contact and when they touch and compress each other a repulsive force, in this case elastic, will begin to be set up until the equilibrium position is established.)

The result of the equilibrium between attractive and repulsive forces is effectively to give the ion a radius (equivalent to the radius of the ball in the above analogy). The distance between the two ions, Na⁺ and Cl⁻ is then the sum of the effective radii. Extensive observation shows that the ionic radius for a given element retains its approximate numerical individuality when the element is in different chemical combinations. In a similar manner the ionic radius characteristic of a complex ion such as NH₄⁺ or CO₃⁻ also retains its approximate numerical value in different compounds. This fact is very helpful when an attempt is being made to analyse a crystal which contains elements or radicles for which the ionic radii have already been established.

In the above ionic bond, which is typical for inorganic crystals, the distance between the constituent atoms is such that each retains its individuality and can be regarded as a separate entity. In contradistinction to this in the homopolar, or valency, bond, the atoms are so close that their electron systems overlap into each other. As a result of this some of the electrons are in effect shared by both systems, and this sharing of common electrons leads to the bonding. The simplest of such types of valency bond is the formation of the hydrogen molecule from two identical hydrogen atoms. The two atoms approach until the electron of each is under the mutual electrical fields of both the nuclei. This situation only arises when the approach is very close, much closer than that required to set up an ionic bond. This electron sharing is equivalent to the exchange of electrons, and it is considered that a rapid interchange is taking place. The bonds of this type are therefore described in terms of "exchange forces." They are specific in direction and are restricted to the number of unpaired electrons. This latter fact explains the simple integral rules of valency.

The crystal structures of metals

The study of metallic structures is a very wide field, for all metals are crystalline and since large numbers of alloys can be produced great complexity can result. In metals a special third type of bonding is found, the metallic bond. A simple metal or alloy crystalline aggregate is a complex built up out of positive metallic ions and free electrons. The arrangement of the positive ions depends upon the particular alloy. Metallic

crystalline states result from two opposing tendencies: (a) the metallic character in which the bonds are non-directional, and (b) specific bonding of each atom to certain of its neighbours, this being to some extent homopolar in character.

The first tendency, that of non-directional forces, leads to simple packing of the atoms into regular arrays, similar to the ways of close packing of spheres. The guiding factor in this packing is the tendency to take up a minimum volume. The face-centred cubic and hexagonal close-packed structures are equally tightly packed and another simple structure, the bodycentred cubic, is only slightly more open. It is possible to fit many metallic crystals into these packing schemes, showing that the first tendency predominates.

The structures of alloys are of great theoretical interest and practical importance. The attempt to discuss alloys in terms of ordinary chemical concepts breaks down, for the chemical composition of an alloy phase often varies within wide limits while the physical properties characteristic of the phase remain effectively unchanged, and even if the composition of the phase lies within fairly narrow limits, the "ideal" composition may correspond to atomic proportions which are far from simple. Alloy structures are best considered as based in the first place on patterns of phase sites-e.g. face-centred cubic, hexagonal close-packed, etc.—the phase pattern being determined for a given alloy partly by the sizes of the various atoms which occupy the phase sites, partly by the ratio of the number of valency electrons to the number of atoms in the unit cell (the Hume-Rothery electron-concentration rule), but hardly at all by the purely chemical properties of the atoms. The distribution of the atoms in a given phase structure among the available sites is determined by the equilibrium between the tendency of the structure to assume the most perfectly ordered arrangement, for which the potential energy is a minimum, and the disturbing influence of thermal agitation which tends to produce a completely random ("disordered") arrangement. A highly ordered arrangement is called a superlattice, but it is to be understood that superlattice formation represents a refinement on the fundamental idea of the phase pattern which largely determines the physical properties of the alloy.

The silver-cadmium system may serve as an example to illustrate the general principles stated above. Pure silver has a face-centred cubic structure. For alloys containing up to 42 atomic per cent. cadmium, the phase-pattern is face-centred cubic, cadmium atoms replacing silver atoms at random, with a corresponding small change in the dimensions of the structure; this single-phase alloy is called a primary solid solution. Alloys containing between 42 and 50 per cent. cadmium contain two phases, one the primary solid solution (42 per cent. cadmium) a phase, the other a new phase-structure (B) containing 50 per cent. cadmium, with a body-centred cubic phase pattern and the silver and cadmium atoms arranged in a perfect superlattice. This intermediate phase exists only for compositions very close to 50 per cent. cadmium-i.e. the limits of the phase are very close together—but this is rather a special case and should not be regarded as typical of alloy systems in general. From 50 to 57 per cent. cadmium the alloy consists of a mixture of β and γ phases; from 57 to 65 per cent. cadmium the alloy contains only the intermediate phase y; and so on, single-phase and two-phase alloys alternating, the last alloys being single-phase primary solid solutions of silver in cadmium, with a phase pattern corresponding to that of pure cadmium.

A further factor found to influence the packing arrangements of the constituents is the ratio of the number of the atoms to the valency electrons in a particular compound. This effect was discovered by Hume-Rothery. Thus a complex lattice with 52 participating atoms may be formed when there are in a compound 21 valency electrons per 13 atoms. Typical examples are Ag₅Cd₈, Cu₅Zn₈, and Cu₉Al₄. The ratio of valency electrons to the number of atoms has been found to be a decisive factor in fixing the structure of a metallic alloy.

Organic crystals

The complex nature of most organic molecules has prevented important advances in the study of organic crystals by X-ray methods. At first little more could be done than the confirmation of structures already derived from stereo-chemistry. The fundamental basis of the chemical method of approach

depended largely upon the recognition of the existence of four valency bonds in carbon and upon the classical conception of the benzene ring. W. H. Bragg first applied X-ray methods to organic crystals and succeeded in obtaining an analysis of the structures of naphthalene and anthracene (two and three linked benzene rings respectively). Later and more extended observations proved the benzene ring to be flat, the carbon atoms forming a regular hexagon with sides 1.41 A long. In this, the three bonds from the carbon atoms lie in a plane, inclined at angles of 120° to each other. Thus not only is the existence of the classical benzene ring confirmed, but in addition its dimensions are now known.

Crystal texture

The properties of a crystalline mass depend not only upon the atomic arrangements characteristic of the particular type of crystal, but also upon the more or less accidental microscopic structure, that is, the *mosaic* structure of microcrystalline components out of which such a crystalline mass may be built. The microcrystals may be either completely ordered or else completely disordered and distributed at random. Furthermore it frequently happens that a particular crystal is in a state of mechanical strain, leading thereby to forms of distortion.

The size of the crystalline conglomerate affects the observations too. When the crystal particles used in the powder method are very small the lines obtained become ill-defined and diffuse. From the line widths it is possible to estimate the sizes of the individual crystals. Ultimately, if the size of the crystalline particles is reduced sufficiently, the crystalline structure is replaced by an amorphous structure. However, even in such cases certain atomic arrangements are more probable than others and this leads to a form of texture which produces diffuse diffraction haloes. Such diffraction haloes with X-rays can even be observed with liquids and gases. It is hardly true to say that the liquids exhibit a crystalline structure, but they do in fact exhibit a texture of molecular arrangements and in a sense this is pseudo-crystalline.

Many natural organic fibres are found to diffract X-rays in such a manner as to imply the existence of some ordered

arrangement even if the arrangement is not to be considered as truly crystalline. There exist in these fibres long complex molecules which consist of similar molecular groupings repeated periodically along the length of the fibre. When a group of fibres is oriented so that the individuals are approximately parallel a crystal-like pattern results. By this line of approach information has been derived about the structures of cellulose and rubber fibres.

The applications of the X-ray technique to the study of the solid state are very wide indeed. Compounds which would be destroyed by chemical analysis can be studied intact as the material suffers no change during X-ray examination. It is possible to examine and study very minute quantities of material, and even if the material is mixed with some other type of crystal separation of the two is not necessary. (Clearly with minute quantities such separation would frequently be impossible.) This avoidance of the necessity for separation is often of great practical value. Much light has been thrown upon the structures of alloys and upon the chemical constitution of many complex compounds. Thus as a typical case, it was proved by X-ray analysis that the formula for amphibole, supposed to be CaMg₃(SiO₃)₄, is in reality (OH)₂Ca₂Mg₅(Si₄O₁₁)₂. Allotropic modifications of crystals can be distinguished by the new methods. By measuring the crystal lattice spacings accurately it is even possible to measure the thermal coefficients of expansion of the crystals and this can be done for very small particles when the powder method is employed.

REFERENCE

[&]quot;The Crystalline State," W. L. Bragg. Vol. 1 (1933).

CHAPTER 11

THE WAVE MECHANICS

The wavelength of the electron

Until 1923 the electron was considered to be entirely corpuscular in nature, but in that year de Broglie proposed a new theory of the electron which, when fully developed, has had remarkable success in explaining certain phenomena. De Broglie proposed that a moving particle, whatever its nature, has wave properties associated with it. Just as in optics the simple ray theory can explain lens properties, etc., but fails to account for fine details such as diffraction and interference effects, to explain which a wave theory is needed, so also the ordinary dynamical corpuscular theory of the electron fails when fine details are examined and by analogy a wave theory of the electron has been adopted. De Broglie proposed that the wavelength λ associated with any moving particle of mass m and velocity v is given by $\lambda = h/mv$ where h is Planck's constant.

It has already been noticed in connection with emission of electricity from surfaces that the electron cannot be treated as a simple particle. In the experiments now to be described electrons must be considered to have wave properties, since they exhibit diffraction effects. The apparent contradiction between the wave and particle theory will be discussed later in more detail. For the moment we shall assume here the validity of the de Broglie relationship, for we shall see that only by this means can the experiments on electron diffraction be accounted for. From this equation it is apparent that the greater the velocity acquired by an electron the shorter will be the wavelength of the associated de Broglie wave. Substitution shows that a 150-volt electron, which has a velocity of 7.2×10^8 cm./sec., has a wavelength of 1 angström, and as the electron velocity is proportional to the square root

of the applied voltage, the wavelength for a 15,000-volt electron is, for example, 0.1 angström.

The wave properties of electrons were first experimentally detected in 1927 by Davisson and Germer who succeeded in measuring the de Broglie wavelength for slow electrons by diffraction methods.

The experiments of Davisson and Germer

The apparatus used is shown in a simplified form in Fig. 11.1. Electrons which have been given a known velocity are sent out from an "electron gun" which is a device for creating a focused strong beam of electrons. These are directed in a

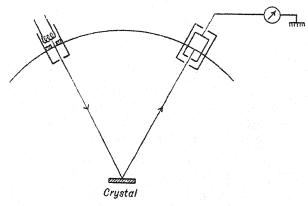


Fig. 11.1

high vacuum at an angle on to the surface of a large single crystal of nickel. Electrons are "reflected" from the crystal planes in different directions, the angular distribution being measured with a Faraday cylinder which can rotate on a divided circle. The Faraday chamber is surrounded by a protecting cylinder to which a retarding potential can be applied, and by varying this the energies of the incoming electrons can be measured as well as the number. Although all the incident electrons strike the crystal with the same velocity, those which come away from the crystal contain amongst them secondary slow electrons excited by collisions with atoms. These can be excluded by making the retarding potential nearly equal to that of the incident electron beam.

It was found from these observations that there is selective reflection, depending upon the velocity of the oncoming electrons. If source and chamber are fixed and the velocity of the electrons is gradually increased, the number reaching the Faraday cylinder follows the curve shown in Fig. 11.2. This selective reflection from the crystal is analogous to the reflection

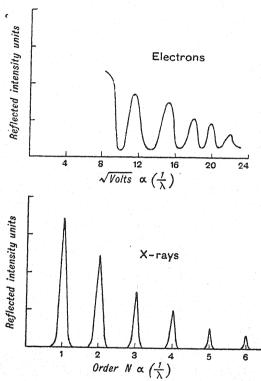


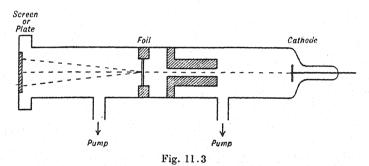
Fig. 11.2 (after Hoag: "Electron Physics")

of X-rays shown by the figure. If the electrons are considered to be simple corpuscles, there is no valid reason for expecting a higher specific reflectivity to be associated with given velocities. However, if it is assumed that the electrons have associated waves, with wavelength varying with velocity in accordance with the de Broglie expression, then exactly the state of affairs which exists when a beam of X-rays of progressively changing wavelength impinges on a crystal is repro-

duced, with consequent selective reflection according to Bragg's law. It is therefore possible to calculate the *effective* electron wavelengths by the application of this law. The values so found are in complete agreement with those given by the de Broglie equation.

The experiments of G. P. Thomson

The experiments described above, which use slow electrons, were followed in 1928 by investigations made by G. P. Thomson with fast electrons. An entirely different experimental method, closely allied to the Debye-Scherrer powder method used in crystallographic analysis, is used. It is illustrated in Fig. 11.3. The right-hand side is a source of cathode rays which can be excited with potentials up to 50,000 volts.



By means of a diaphragm tube a fine beam of electrons is isolated and allowed to fall upon a very thin gold foil. The thickness of the foil is of the order of 10⁻⁶ cm. A special technique is required to produce such extremely fine films. In one method metal is sputtered on to a celluloid acetate base, the latter then being removed by immersion in acetone.

If the electrons are corpuscles, the foil will scatter them on to a fluorescent screen so as to form a regular patch, the intensity of which will fall off away from the centre. The distribution can be observed either directly with the fluorescent screen or using a photographic plate. Instead of a continuous patch the appearance observed is similar to that shown in Plate IIID in which distinct sharp rings can be seen. These rings are closely similar to those found when employing the powder

method for the X-ray analysis of crystals, and have a similar origin. The foil consists of a random distribution of minute metallic crystals and those which are just at the correct angle scatter electrons in accordance with Bragg's law. The circles are therefore due to cones of diffraction, the intersection of a cone with the photographic plate producing a circle. The material of the foil and its state of aggregation decides the form of the pattern which is ultimately dependent upon the crystal structure of the foil material. Clearly this experiment demonstrates in a very striking manner that the incident electrons behave as waves, since diffraction patterns can only be produced by waves. There remains one possibility to be examined, namely, is it possible that the electrons in their passage through the foil generate secondary X-rays which would show diffraction effects on the screen? Thomson eliminated this possibility in a very simple way. When a magnet is brought up to the discharge tube the electron beam is deflected and with it the complete pattern, proving that the latter is due to electron waves and not due to X-rays, which would not be deflected.

From the diffraction rings the wavelength can be easily calculated and is found to be independent of the foil material, being determined only by the velocity of the electrons. Measurement showed that the observed wavelengths are those given by the de Broglie equation, to within the 2 per cent. experimental error. The phenomenon described above is called electron diffraction and has recently become a powerful method of crystal analysis. Diffraction patterns can also be produced by reflecting electrons from crystal surfaces at small angles. As in the case of transmission, the different diffraction rings are due to reflections from different crystal planes, Surface films and surface crystalline structure can be studied by analysing the electron diffraction rings obtained by reflection.

The dualistic nature of matter and waves

It is clear from the experiments just described that the electron at times must be treated as a wave. On the other hand, experiments such as those in which e/m is measured

show that the electron must in these cases be considered to be a particle. In 1932 Stern proved that a beam of atoms or molecules, when "reflected" from a crystal surface, exhibits diffraction effects exactly as does a beam of electrons. The wavelengths of the associated "matter waves" exactly fit into the de Broglie expression, which therefore holds for all matter. The characteristic wave properties of electrons must also be attributed to all matter, electrons, protons, molecules, etc. This fundamental fact leads to a new form of mechanics, "wave mechanics," which replaces classical dynamics when the fine structure details of matter are to be considered.

The extension of a wave theory to matter has its counterpart in the treatment of light waves. Although light unquestionably exhibits the well-known phenomena of electromagnetic waves, yet at the same time it has a true particle aspect, photons or light quanta having a definite momentum, as proved by the existence of the Compton effect. The only conclusion that can be reached is that both matter and radiation have a dual "wave-particle" nature.

The first attempt to explain this dual nature was made by Schroedinger by treating each electron as a wave packet. Consider, for example, two sound waves with different wavelengths, but travelling with the same velocity. At certain points in the path the amplitudes add up to maxima and minima producing beats. These successively move past any point with the same velocity as the waves. If we consider two sound waves with different velocities (i.e. in a dispersive medium) the velocity of the maxima and minima is not that of the component waves. If a large number of wave trains, each with slightly different velocity and wavelength, are combined together, these can be selected so that the vibrations cancel except over a very small region in space, where they add up to form what is called a wave packet. Such a packet moves forward with a velocity of its own, the "group velocity." The individual waves forming the packet may be considered to possess an average velocity, the "phase velocity."

It can be shown that a wave packet must ultimately dissipate. If the electron is considered to be a wave packet, it is necessary to postulate the existence of a "guiding wave" for it. Without such an assumption the wave packet theory is untenable. This guiding wave is that described by Schroedinger's equation, which will be discussed later. Physically this equation related the amplitude of the guiding wave to the probability of finding an electron at a point. If the amplitude of the guiding wave is at any point zero, there is only an infinitesimal probability of finding the electron at that point. The mechanical process has associated with it a wave process the square of whose amplitude gives a measure of the probability of the event taking place at the point considered. The value of the amplitude is therefore vanishingly small at all points except in the wave packet, hence one can expect to find an equivalent particle within the range of the packet and thus effectively it travels with the group velocity v. Thus the wave packet plus guiding wave have the properties of a particle moving with velocity v and also exhibit diffraction and interference effects. So the difficulty is resolved. The particle and wave properties are strictly complementary. If we devise an experiment showing wave properties we are debarred from seeing particle properties, and vice versa. An electron cannot at the same time behave as a particle and a wave, it acts as one or the other in any experiment that can be devised.

The wave and particle properties can be equated according to the following scheme.

Wave property— Wavelength Frequency

Particle property— Momentum Energy

The only link between the two concepts is h, in the de Broglie expression.

Heisenberg's uncertainty principle

A principle of far-reaching importance was proposed by Heisenberg in 1927. It was arrived at from considerations of the dualistic nature of the electron. The very definition of a corpuscle implies that at any instant it has a definite momentum and occupies a definite position in space. Unless we can simultaneously determine both the momentum and the position in space we cannot actually say that a "particle" in the accepted sense has been observed. Since the electron is somewhere within the wave packet, moving with group velocity,

uncertainty about the defined particle velocity arises, for the group is not infinitely narrow and has a velocity spread. It is impossible to know where, within the group, the electron actually is and what is its exact velocity. For a long wave packet, with many crests, the position of the electron is very uncertain but the velocity spread is very small so that the particle velocity is fairly accurately known. In a short packet the position of the particle is more or less fixed, but as the velocity spread of such a packet can be shown to be large, the particle velocity is indeterminate. Either the position or the velocity can be known accurately, but not both, and one has a doubtful value. It can be proved that it is impossible to determine simultaneously both the momentum and the position of a particle with accuracy. This is shown by the following treatment.

A wave packet representing a particle has a finite length Δx and the extreme ends have a wave number difference $\Delta \nu$. It can be shown from general wave theory $\Delta x \sim 1/\Delta \nu$, the length of the wave packet being inversely proportional to the difference in wave numbers of the two ends of the packet. Thus $\Delta x.\Delta v \sim 1$. Since we have also $\lambda = h/mv = h/p$ where the momentum mv is written as p, then

 $p = h\nu$

differentiating gives

 $\Delta p/h = \Delta v$

when, by substitution we get

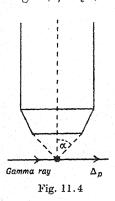
 $\Delta x. \Delta p \sim h.$

It is clear that Δx represents the uncertainty in determining the exact location of the particle within the packet, and as there is a range of wave numbers Δv in the packet Δp is the uncertainty in evaluating the momentum. This then is Heisenberg's uncertainty relationship which states that the product of the uncertainties in determining position and momentum is approximately equal to Planck's constant h.

The more exactly we define the position, i.e. the smaller we make Δx , the larger becomes Δp and vice versa. The relationship shows that it is impossible to measure both the position and the momentum accurately simultaneously. Clearly this is fundamental since it sets a limit to the possible accuracy of observation. The uncertainty arises from the smeared-out

properties of wave packets compared with finite points representing the centres of corpuscles.

No matter what method is tried it is impossible to avoid the consequences of the uncertainty relationship. Consider, for example, an attempt to observe accurately the position of a particle with a microscope of extremely high resolving power. Optical theory shows that the limit of resolution depends upon the wavelength of the light employed to view the particle. Let us imagine that we can use the shortest available wavelengths, γ -rays, achieving by this means the highest possible



resolution in determining the position of the particles. Let ν be the frequency of this illuminating radiation, and let the particle under observation be an electron. From classical optical theory the position of the particle is known to within an amount Δx where $\Delta x \sim \lambda/\sin \alpha$, a being the angular aperture of the microscope.

In order to see the electron, the radiation reaching the observer must of necessity have been first scattered from the electron, which, in the scatter process, suffers a recoil due to the Compton effect, the

momentum change being of the order of magnitude $h\nu/c$ in an indeterminate direction. The uncertainty in the direction of the scattered light quantum is just the same as the uncertainty in the recoil direction of the electron. As the experimental conditions are such that the particle is actually seen, the quantum of scattered radiation certainly enters the microscope. The uncertainty therefore in the direction taken by the quantum of radiation is equal to the angular aperture of the microscope. Thus perpendicular to the axis of the instrument the component of momentum is uncertain to the amount $\Delta p \sim (h\nu/c) \sin a$.

Multiplying, we have

 Δp . $\Delta x \sim (h\nu/c) \sin a$, $\lambda/\sin a$

and as we get

 $\lambda \nu = c$ $\Delta p \cdot \Delta x \sim h.$

Thus once again there is no escape from the uncertainty relation, and in fact whatever attempt is made, the conclusion is forced upon us that it cannot be overcome. The philosophical and scientific implications of Heisenberg's principle are very wide indeed as it is necessary to abandon the law of exact causality in the classical sense. In effect, probability takes the place of exactness in physical science. Phenomena are to be described only in terms of probability and statistical distributions. Events impossible to the classical theory are found, when treated by wave mechanics, to have a very small but finite probability of taking place.

The Schroedinger wave equation

The guiding wave obeying the Schroedinger wave equation has been referred to as necessary to explain the non-dissipation of the electron wave packet. In this section the derivation of the wave equation will be considered. Schroedinger begins by considering the dynamical equation for the propagation of elastic waves. Assuming for the moment propagation takes place in the x direction only, this is

$$\frac{\partial^2 \Psi}{\partial x^2} = \frac{1}{c^2} \frac{\partial^2 \Psi}{\partial t^2}$$

 Ψ being the wave displacement and c the wave velocity. The solution to this equation gives Ψ as a periodic displacement in terms of time, namely, $\Psi = \psi e^{iwt}$. In this, ψ is a function of x but not of t, and w is equal to $2\pi\nu$ where ν is the frequency. On differentiating this twice with respect to x and t we get

$$\begin{split} &\frac{\partial^2 \Psi}{\partial t^2} = -w^2 \psi e^{iwt} \\ &\frac{\partial^2 \Psi}{\partial x^2} = \frac{\partial^2 \psi}{\partial x^2} e^{iwt} \end{split}$$

By substitution in the equation for elastic waves we get

$$\frac{\partial^2 \psi}{\partial x^2} e^{iwt} = \frac{-w^2}{c^2} \psi e^{iwt}$$
$$\frac{\partial^2 \psi}{\partial x^2} + \frac{w^2}{c^2} \psi = 0.$$

i.e.

Since $\lambda = \frac{c}{v} = \frac{2\pi c}{w}$ this can be rewritten

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{4\pi^2}{\lambda^2} \psi = 0.$$

So far the treatment is general. The wave mechanics concept is now introduced by substituting $\lambda = h/mv$ from the de Broglie equation leading to

 $\frac{\partial^2 \psi}{\partial x^2} + \frac{4\pi^2 m^2 v^2}{h^2} \psi = 0.$

In this equation ψ is the amplitude of the wave associated with the moving particle.

If W is the total energy, V the potential energy, and $mv^2/2$ the kinetic energy of the particle, then $mv^2/2 = W - V$, and if this is substituted in the wave equation just derived above

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{8\pi^2 m}{h^2} (\mathbf{W} - \mathbf{V}) \psi = 0.$$

If this is generalised to include wave motion for the x, y, and z axes it takes on the form

$$\nabla^2 \psi + \frac{8\pi^2 m}{h^2} (\mathbf{W} - \mathbf{V}) \psi = 0.$$

The Laplacian, $\nabla^2 \psi$ is

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + \frac{\partial^2 \psi}{\partial z^2}.$$

The above relationship is Schroedinger's wave equation.

The wave picture of the hydrogen atom

It is necessary to replace the orbital theory of Bohr by an equation of the Schroedinger type in order to arrive at a more correct picture of the hydrogen atom. Actually the wave mechanics aspect is closely allied to the simpler orbital theory. Instead of considering an electron moving in an orbit the formula implies that there is a definite probability of finding the electron in a given volume dv. This probability is $\psi\psi^*dv$ where ψ^* is the complex conjugate of ψ .

The potential energy V for an electron distant r from a

nucleus with charge Ze is $-e^2\mathbf{Z}/r$ and, making this substitution, the Schroedinger equation of such a system becomes

$$\nabla^2 \psi + \frac{8\pi^2 m}{h^2} (W + e^2 Z/r) \psi = 0.$$

To solve this expression solutions must satisfy boundary conditions such that ψ and its first derivative are everywhere single valued, continuous, and finite. Such solutions are called "eigen-functions," the values so found for W being of great importance. It is found that the Schroedinger equation can only be solved when $W=2\pi^2me^4Z^2/n^2h^2$ where n has integral values 1, 2, 3, etc. For all other values of n the de Broglie waves in the Coulomb field of the nucleus destroy one another by interference.

The remarkable fact is that these possible energy values are exactly the same as those given by Bohr's orbital theory. Thus the experimentally observed Bohr energy levels are predicted from the wave mechanics without the introduction of any of the arbitrary assumptions made by Bohr. Furthermore, many observations not in accord with the Bohr theory can be explained with the aid of the newer theory. To a certain extent the Bohr theory can be looked upon as a first approximation. More detailed analysis of the equations shows that the quantum numbers m and l also appear in eigen-function solutions.

Electrical charge distribution for atomic states

The quantity $\psi\psi^*dv$ is the probability of finding an electron in the volume element dv. When $\psi\psi^*$ is plotted as a function of r the distance from the nucleus the resulting curve (for the hydrogen atom) is that shown in Fig. 11.5a. The electric distribution is spherically symmetrical. The probability of finding the electron within a given volume element is large near to the nucleus and falls off rapidly with increasing r. The electron distribution function, defined as $D=4\pi r^2\psi\psi^*$, is shown in Fig. 11.5b. The quantity Ddr is therefore the probability of finding the electron in the shell bounded by spheres of radii r and r+dr. D is a maximum at the radial distance r=a where a is the radius of the first Bohr orbit.

The Bohr orbit is therefore replaced by the electron distribution function. To a first approximation the electron is situated at the Bohr orbital distance but it is not an effective point in space. It is to be regarded as being "smeared out" over

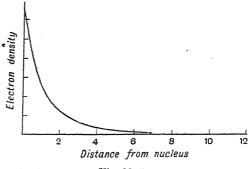
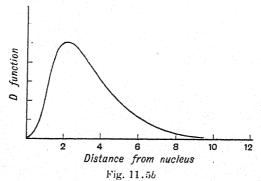


Fig. 11.5a

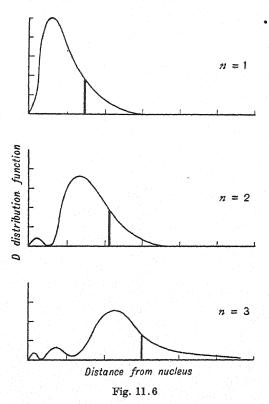
the range of the D curve. The calculated electron distributions for the quantum states with l=0 and n=1, 2, 3, are shown in Fig. 11.6 (the horizontal scales of the figures are different; unit interval is proportional to n in each case). The vertical line in each curve represents the amplitude of the corresponding



(purely radial) motion (k=0) derived from the dynamical theory. Although the D function theoretically extends to infinity it will be seen that its value is virtually zero except within a radius of the order of magnitude of the major semi-axis of the ellipse of the older orbital theory.

The older theory can only give approximately correct values for integrated or average fields, etc., and it is found that the values derived from it are in closer agreement with those

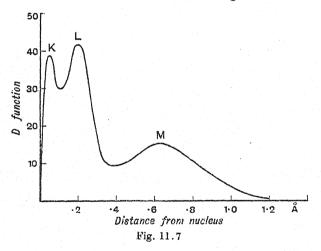
given by wave mechanics methods if k is replaced by $\sqrt{l(l+1)}$ instead of by (l+1). Even then certain equations require modification. For example orbital formulæ involving k^3 when derived more precisely by wave mechanics methods involve instead the quantity $l(l+\frac{1}{2})(l+1)$. These modifications were actually suggested empirically before being derived



theoretically. The theory has justified the application of these empirical changes which were originally adopted to fit observation. For descriptive purposes and also for practical application in experimental work, the older model is still of immense value, but for the interpretation of very refined observations and for the derivation of complex formulæ, such as those for atomic and nuclear magnetic moments, the wave mechanics improvements require to be introduced.

Electron distributions for many electron atoms

The methods first applied with success to the hydrogen spectrum have now been extended to atoms of higher atomic number. Hartree has developed a method whereby the distribution for an atom with a number of closed shells and with outer valency electrons can be deduced. The alkali atoms, for example, have one outer valency electron and removal of this leaves an alkali ion which has the same closed shell structure as a rare gas atom. The total electron probability function for a closed group depends only upon r, hence alkali ions have a spherically symmetrical distribution. A typical result is that shown for K^+ in Fig. 11.7. There are



three maxima in the distribution, and these can be identified with the K, L, M shells respectively in the orbital model. The wave mechanics analysis thus also leads to a structure effectively equivalent to the existence of closed shells.

REFERENCES

[&]quot;Atomic Spectra and Atomic Structure." G. Herzberg. (1937.)
"Atomic Physics." M. Born. (1935.)

CHAPTER 12

ELECTRON COLLISIONS IN GASES.

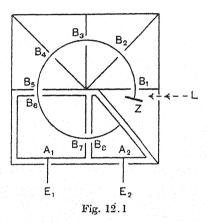
The Ramsauer effect

On sending a beam of electrons through a gas, collisions with gas atoms take place. These can be divided into two kinds: (a) elastic collisions, in which the electrons are merely deflected from their paths without loss of energy, and (b) inelastic collisions, in which there is an energy interchange between the electron and the atom. The total kinetic energy is conserved in the first case, whilst in the second type of collision, energy of motion may be, and usually is, converted into radiation. Some of the electrons in the incident beam are effectively absorbed by colliding with gas atoms, the number so removed from the beam defining the cross-section of the atoms for the process in question.

Consider a beam of electrons, with a cross-section of 1 sq. cm., entering a given gas. Let A be the effective cross-section of each atom of the gas. If there are n atoms per cubic centimetre the total effective absorbing cross-section in a path dx cms. long is nAdx. If the cross-section of the electron beam contains N electrons at the point x in its path, the number absorbed, -dN, is equal to -(nAdx)N. Integrating this expression gives $N = N_0 exp(-nAx)$, N_0 being the number of electrons in the beam at x=0. This is the ordinary expression for logarithmic absorption, hence by measuring the reduction in the number of electrons in a beam brought about by the passage through a thickness of the gas x cms., the absorption coefficient, nA, can be determined. The number of atoms per cubic centimetre, n, is known from the gas pressure, thus the effective atomic cross-section A can be derived. Clearly nA is inversely proportional to the mean free path of the electrons. Kinetic theory shows that $nA = 1/\lambda$ where λ is the mean free path. On the assumption that both

atoms and electrons are hard spheres, A should be quite independent of the velocity of the incident electrons and equal to the value, deduced from measurements of viscosity, etc., for collisions between molecule and molecule in the same gas.

The absorption of electrons has been studied by Ramsauer with the apparatus shown in Fig. 12.1. A beam of ultraviolet light L liberates photo-electrons from the plate Z. These are accelerated to the slit B_1 by an electric field. A magnetic field is applied perpendicular to the plane of the paper, deflecting the electrons into circular paths. The system of slits, B_1 ... B_8 enables a beam of electrons of constant velocity to be selected, since the slits lie upon a

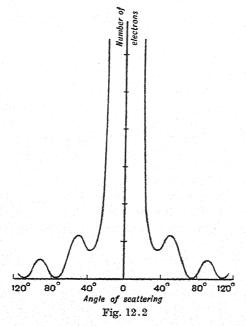


circle. A_1 and A_2 are Faraday cages connected to electrometers by E_1 and E_2 . The apparatus is filled with gas at pressures that can be varied. The electrons scattered in the path from B_6 to B_7 are measured by E_1 and those passing beyond this by E_2 . Hence N_0 is the sum of the currents to E_1 and E_2 since this is the total beam strength, and the current to E_2 is N, the intensity of the beam after passing through the distance B_6 — B_7 . The slit system keeps away from the Faraday cages those electrons whose velocities have been affected by collisions.

Ramsauer did not find A to be constant, particularly when the rare gases were used as the absorbing atoms. The effective cross-section of the atoms was found to depend upon the velocity of the electrons. When the electron velocities are high, the observed values approach those calculated from the kinetic theory of gases, but as the velocities decrease A rises to a maximum and then falls sharply to a value below that given by theory. The existence of these anomalously small cross-sections prove that atoms are not solid spheres but have an open structure.

Elastic scattering of electrons

The absorption experiments of Ramsauer throw some light upon the nature of the gas atoms producing the absorption.



When, however, the elastically scattered electrons are studied in more detail light is also thrown upon the nature of the electron. An electron beam is sent into a rarefied gas (e.g. mercury vapour) and only those electrons which have not suffered any energy loss at all are selected and studied. The angular distribution of the scattered electrons is of particular interest. Most of the electrons are scattered through small angles, but as the scattering angle increases the number

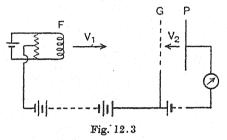
scattered goes through maxima and minima. This is shown

in Fig. 12.2.

The shape of the scattering curve is closely analogous to the halo diffraction pattern made by sending a beam of light through a suspension of fine particles, or through a fine powder spread upon a glass base. These observations confirm that the electron has wave properties associated with it, the maxima and minima being due to diffraction of electrons.

Critical potentials

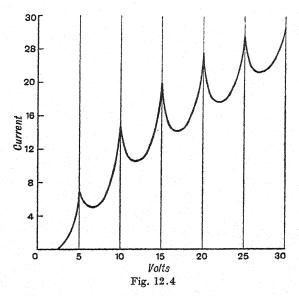
We shall now consider what takes place when there are inelastic collisions between electrons and gas atoms. In such collisions the electrons part with some of their energy and this



does not reappear as kinetic energy of struck atoms but is emitted in the form of radiation. Such inelastic processes only take place when the electron energy exceeds a certain critical value depending upon the gas through which the electrons are passing. The experiments of Franck and Hertz in 1914 first demonstrated the existence of inelastic collision. The method of investigation used by these authors is shown in Fig. 12.3. The apparatus contains a trace of the gas to be studied. Electrons, generated by the hot filament F, are accelerated to a grid G by a potential V_1 . A much smaller opposing potential V_2 is applied between the grid G and a plate G. The gas pressure and dimensions are such that the distance G is much less than the mean free path for the electrons whilst the distance F is slightly greater than the mean free path.

As the applied voltage V_1 is increased from zero, the current reaching P rises, following the curve shown in Fig. 12.4. The

first peak, at voltage V, means that a number of electrons with a critical energy $mv^2/2 = Ve$ begin to lose nearly all of this energy by inelastic collision before reaching G. They are therefore turned back from P by the small retarding potential V_2 , the result being that the number recorded falls after the voltage V instead of increasing steadily. The electrons at first only acquire sufficient velocity when they are already close to G, but as V_1 is increased the point at which the critical velocity is reached moves back towards F so that more and more electrons suffer inelastic collision and the



current falls until V_1 becomes equal to $V+V_2$, after which there is again an increase.

Another peak appears at twice the critical potential, since now electrons can suffer two inelastic collisions before reaching G. This effect is repeated if the voltage V_1 is still further increased, and in fact it is found that the peaks are all separated by the same voltage, V. This is termed an excitation potential. More refined observation shows that each type of atom has associated with it a number of characteristic excitation potentials.

The observations prove that energy can only be absorbed in

discrete quanta. The amount of energy absorbed is that required to lift an electron from the normal, or ground state, to a higher orbit. This is therefore a direct proof of the existence of distinct energy levels within an atom. When the incident electrons have a sufficiently high velocity, ionisation by collision sets in so that by this experimental procedure ionisation voltages can be measured.

Controlled excitation of spectra

If the quantum theory of spectra is true, it is clear that no radiation can be emitted by an atom unless it receives energy at least equal to that necessary to raise the electron from the lowest to the second lowest orbit. Corresponding to this energy is the first critical potential, or resonance potential of the atom, \overline{V} . It follows that if v is the velocity of the electron required to do this, we have $mv^2/2 = Ve = h\nu$ where ν is the frequency of the resonance line. If a gas is bombarded with electrons with just this energy, the optical electron in the atoms can be raised to the first excited level and after a short time (of the order of 10-8 sec.) falls back to the normal state, a spectrum line being emitted in the process. Only a single line will be emitted. When the velocities of the incident electrons are increased a second line can be emitted. With still higher velocities a third line, then a fourth, etc., will in turn be radiated, until, when the ionisation potential is reached, the whole arc spectrum will be given out.

This prediction of the quantum theory was first verified by Franck and Hertz who showed that with mercury vapour, an impacting voltage of 4.9 volts resulted in the emission of the line 2537 A, the resonance line of mercury. Substituting for this wavelength in the expression $Ve=h\nu$ gives V=4.87 volts, a value in excellent agreement with that observed.

The phenomenon of single line excitation is illustrated by the photographs in Plate IIc taken with magnesium vapour. At an impact velocity of 3·2 volts the single principal resonance line is emitted. At 6·5 volts a second line is radiated, and as the voltage is increased the individual lines of the spectrum appear one by one.

A knowledge of ionising potentials is of particular value to

both spectroscopy and chemistry. They can be obtained either from observations of the limits of line series, the study of the spectra of dissociating molecules, or by direct electron impact experiments. There is excellent agreement between the values derived by the different methods. Ionisation potentials are found to vary from about 4 to 24 volts for normal spectra, but for multiply ionised atoms as much as 390 volts may be required.

REFERENCE

"Atoms, Molecules and Quanta." A. E. Ruark and H. C. Urey. (1930.)

CHAPTER 13

RADIOACTIVITY AND THE RADIOACTIVE TRANSFORMATIONS

Historical

The discovery of radioactivity in 1896 by Becquerel proved to be the beginning of one of the most fruitful and important developments of modern physics. Becquerel originally noticed that a uranium salt which was wrapped up in paper, and had been so for a long time, emitted penetrating radiations which affected a photographic plate. These radiations were given off quite spontaneously, requiring no previous excitation of the uranium salt in order to cause them to be emitted. photographic action was extremely weak, but investigation proved that in addition to this power, the radiations were, like X-rays, also able to produce ionisation in the gases through which they passed. Rutherford then began a series of investigations of this phenomenon. He first proved that the radiations given off by the uranium salt were of two distinct kinds. The one, a-radiation, was easily absorbed by thin sheets of matter and was capable of producing intense ionisation. other, \beta-radiation, was found to be much more penetrating than the former and correspondingly much less effective as an ionising agent. Villard later discovered that a third still more penetrating radiation, γ-rays, was also emitted. This property of emitting radiations is an atomic property being entirely independent of the state of chemical combination of the uranium. It is called radioactivity.

After the discovery of this phenomenon, the question immediately arose whether radioactivity was an atomic property characteristic of uranium only and to test this point an extensive search was carried out. As a first result thorium was found to possess radioactivity comparable in intensity with

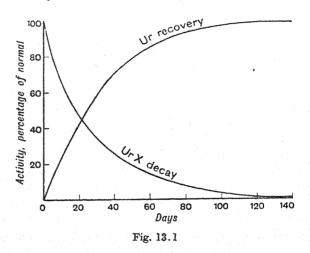
that of uranium and similar in nature, whilst potassium and rubidium were found to exhibit a relatively weak β-ray activity, the intensity of the latter being approximately onethousandth that of uranium. A very important step was taken in 1898 by Mme. Curie who noticed that pitchblende, a uranium mineral, was four times more active, weight for weight, than a pure uranium salt itself. By fractional crystallisation and the use of an electroscope to indicate the concentrations of radioactive material, Mme. Curie succeeded in separating from the mineral two new extremely active materials of high atomic weight which were named polonium (Po) and radium (Ra). The activity of these materials is very great, that of radium being several millions of times that of uranium, for equal weights.

The natural production of radioactive matter

In 1900 Crookes showed that if uranium salt is precipitated from solution by ammonium carbonate and then redissolved in excess of the reagent, a residue is left which possesses the whole of the radioactivity, the original uranium being apparently deprived of all radiating power. This active residue he called uranium X (UX) and its occurrence appeared at first to contradict the view that radioactivity is an atomic property of uranium. However, Becquerel found that if the uranium and the uranium X were laid aside separately for a year, a remarkable occurrence had taken place. The uranium entirely recovered its former radiating power, whilst the uranium X was, as far as could be observed, inactive. Careful investigation showed that the uranium X activity decayed exponentially, whilst the radiations from the uranium recovered in a complementary manner. The decay and recovery curves are shown in Fig. 13.1.

The figure shows that after twenty-four days the activity of the uranium X has fallen to half its initial value. The sum of the activities of the uranium and the uranium X remains constant. This explains why the activity of the unseparated mixture is constant. The experimental decay curve for the uranium X obeys the law $I_t = I_0 exp(-\lambda t)$ where I_t is the intensity of the activity after time t, Io the original intensity, and λ a constant, called the transformation constant since it

is a measure of the rate of transformation of the active material. In a similar manner the recovery curve of the uranium is found to fit the law $I_t = I_0 - I_0 exp(-\lambda t)$, the symmetry of the curves showing that the same constant λ is involved in both decay and recovery. The time required for the decaying intensity to reach half the original value is called the *period*. All radioactive bodies exhibit decay, but the periods in some cases are so long (reaching 10^{10} years) that these may be considered to have a constant activity. On the other hand, periods as small as 10^{-7} second have also been



inferred. Decay periods covering the whole of this wide range have been found in different radioactive bodies, a large number of which are now known. It has been found quite impossible to effect any alterations in the rate of decay or recovery of any radioactive body by physical or chemical means. The rate of production of active substances like uranium X is a natural phenomenon, a property of uranium, that cannot be altered.

Radioactive decay

The decay and recovery, although quite independent phenomena, take place at exactly the same rate. This can only be explained if it be assumed that there is a constant rate

of production of, for example, uranium X by the parent radioactive body. The activity of the newly formed matter must decrease exponentially immediately from the time it is formed. The new matter is, in this case, chemically different from the parent which has produced it. As we have seen, mathematically the decay law is of the form $I_t = I_0 exp(-\lambda t)$ which can be rewritten as $N_t = N_0 exp(-\lambda t)$, where N_t is the number of atoms as yet unchanged after time t, No being the number at the beginning. Differentiating this expression gives $dN_t/dt = -\lambda N_t$, i.e. the rate of change is proportional to N_t the number as yet unchanged. Consider the growth of uranium X in uranium. If q_0 particles of uranium X are produced per second by a given mass of uranium, and N is the number of particles of uranium X present at time t after the complete removal of the initial amount of uranium X, the following holds. As the rate of change of uranium X atoms due to decay is $-\lambda N$, the net rate of increase is

$$dN/dt = q_0 - \lambda N$$

which integrates to $N=Aexp(-\lambda t)+B$. As N=0 when t=0and as N reaches a steady value N_0 when t is infinite

$$N = N_0 - N_0 exp(-\lambda t)$$
.

This value of N is the number of particles of uranium X present at time t after the removal of all the uranium X originally The expression for N therefore gives the law of recovery. This agrees with the experimental values shown in Fig. 13.1.

The theory of disintegration

The phenomenon of the spontaneous production of radioactive matter was first explained by Rutherford and Soddy who showed that the atoms of the radioactive elements must undergo disintegration. In disintegrating, a radioactive atom sends out α or β -radiations which are material in nature, that is, they consist of particles. According to present views, we should say that a small definite portion of the nucleus of the radioactive atom is ejected with very high velocity and a different atom is left behind. The residual atom is also radioactive, leading to a long chain of radioactive atoms in the

form of a series. The sequence goes on until a stable substance is reached. For both the uranium and thorium series of elements the stable end-product is an isotope of lead. The transformation constant of a particular substance is interpreted as the probability per unit time that any atom of that substance will disintegrate; this probability is, of course, the same for all atoms of the substance in question.

The deduced probability of atomic disintegration is not high. For instance, in thorium about one atom in 10^{18} breaks up in a second. The initial disintegration of the thorium or uranium results in the ejection of an a-particle, only one being ejected from each disintegrating atom. It will be shown later that the a-particle has a mass four times that of the hydrogen atom, the residual atomic weight after the a-particle disintegration being four units less than that of the parent atom. In some disintegrations only a β -particle is ejected, and as a β -particle has a charge but only very small mass, the nuclear charge only and not the nuclear mass is effectively changed by the disintegration.

Radioactive equilibrium

If one begins with a pure sample of radioactive matter, after an interval it will have associated with it various radioactive products. After some time an equilibrium position will set in wherein the rate of creation of new material will be balanced by the rate of decay. The equilibria which can result in various particular cases can be studied by a method due to Rutherford. When a large number of products are possible the calculation becomes complex. As an example of the method we shall consider the equilibrium which is set up when a radioactive body A can disintegrate into a radioactive body B which again can disintegrate into a radioactive body C. The transformation constants for these three bodies are not necessarily the same; let them be λ_1 , λ_2 , λ_3 respectively. Let P, Q, R, be the number of atoms of A, B, C, present in the mixture after time t. The rate of increase of the B atoms is the difference between the number supplied by A and those of B naturally disintegrating. Therefore

$$dQ/dt = \lambda_1 P - \lambda_2 Q$$
.

RADIOACTIVITY AND THE RADIOACTIVE TRANSFORMATIONS 21

In a similar manner the rate of increase of C atoms is given by

$$d\mathbf{R}/dt = \lambda_2 \mathbf{Q} - \lambda_3 \mathbf{R}$$
.

If n is the number of A atoms present at the beginning, i.e. at the time when B and C are not present, we know that $P = ne^{-\lambda_1 t}$ from the form of the elementary decay curve. Substituting gives

$$dQ/dt = \lambda_1 n e^{-\lambda_1 t} - \lambda_2 Q$$

which integrates to

$$Q = n(ae^{-\lambda_1 t} + be^{-\lambda_2 t}).$$

By substitution

$$.a = +\lambda_1/(\lambda_2 - \lambda_1),$$

and as Q=0 when t=0 then also

$$b = -\lambda_1/(\lambda_2 - \lambda_1) ;$$

therefore

$$Q = \frac{n\lambda_1}{\lambda_2 - \lambda_1} (e^{-\lambda_1 t} - e^{-\lambda_2 t})$$

By substituting for Q in the expression for R we get

$$R = n(pe^{-\lambda_1 t} + qe^{-\lambda_2 t} + re^{-\lambda_3 t})$$

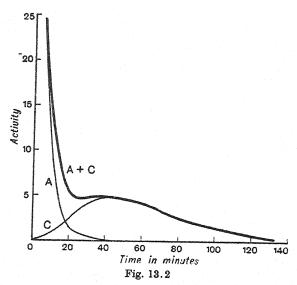
where

$$p = \frac{\lambda_1 \lambda_2}{(\lambda_1 - \lambda_2)(\lambda_1 - \lambda_3)} \cdot q = \frac{-\lambda_1 \lambda_2}{(\lambda_1 - \lambda_2)(\lambda_2 - \lambda_3)} r = \frac{\lambda_1 \lambda_2}{(\lambda_1 - \lambda_3)(\lambda_2 - \lambda_3)}$$

If the constants λ_1 , λ_2 , λ_3 are known, the expressions for P, Q, R give the quantities of A, B, C, present at any time t.

These formulæ represent a special case involving three products only. The general expression for n products is much more complex. The formulæ are best applied graphically. Fig. 13.2 is an example. If a plate be held for a short time in the neighbourhood of radium it coats with a deposit of radium A. If the plate is removed and its a activity measured over a period of some $2\frac{1}{2}$ hours, the intensity of the radiation is found to fall off in the manner shown by the heavy curve marked A+C. Radium A disintegrates with a period of 3.05 minutes (with a disintegration) into radium B which is a active, decaying into radium C which is a-ray active, the period being 19.7 minutes. The decay of the radiation from A is shown by the falling curve A and the growth of the radiation from C by the curve C. The heavy observed curve is the resultant of these two. In practice the curves A and C are

deduced from the combined curve and from these the separate values of the decay periods are calculated. Considerable skill is required in interpreting curves built up out of many components. The curves of the separate components have in a



large number of cases been verified by separating the individuals and making a separate decay determination with each.

The radioactive emanations

An interesting characteristic of three of the known radioactive bodies is that they continuously emit an emanation which behaves like a radioactive gas of high atomic weight. Over 40 radioactive bodies fit into the radioactive series, but of these only radium, thorium X and actinium X give off emanations. Since radium emanation is much more striking in its effects than the other two, its properties only will be considered. The properties of the others are very similar.

Rutherford first proved by a few simple experiments that radium emanation was a true gas which could be transported away from the parent by a current of air. It passes through cotton wool and bubbles through water without losing any of its activity and behaves thus differently to the gaseous ions

produced in a gas by any means. The emanations lose their activity rapidly, decaying exponentially like any other radio-active body, the periods for the emanations from radium, thorium X, and actinium X being 3.8 days, 54 seconds, and 3.9 seconds respectively. The decay of the radium emanation has been measured directly by isolating it from its parent, radium. For reasons which will appear later the emanations from radium, thorium X, and actinium X are called radon, thoron, and actinon.

Some properties of radon

The rate of production of radon by radium is constant and not affected by physical or chemical agencies. The same is true for the rate of decay of the radon itself. Although the radon is generated at a fixed rate by radium, it can be occluded in the parent compound (say radium bromide) and is therefore released when the latter is dissolved or fused. The radiation from radon consists only of α -particles, but when kept some time β -particles and γ -rays also appear. These do not, however, come from the radon itself but from the active deposits left when the radon disintegrates.

A direct chemical atomic weight determination is difficult since the quantities of radon available are very small. An approximate estimate has been made by comparing the rate of diffusion with that of mercury vapour. From its properties radon can be classed as a chemically inert gas of the same family as helium, neon, argon, krypton, xenon, and the same is true for thoron and actinon. It is for this reason that the radioactive emanations have been separately named with "on" endings. A conclusive proof of the fact that radon is a true gas and not an aggregation of particles was given by Rutherford and Soddy who were able to condense it at temperatures below -150° C. Like many other gases radon can be absorbed by water and porous substances.

When radon disintegrates it loses its gaseous nature and the residual changed atoms are deposited as solid active deposits upon any available surface. The amount of active deposit formed is proportional to the amount of radon present. Being

active it also disintegrates, giving rise to a long chain of products. The volume of radon which is in equilibrium with a fixed quantity of radium is constant, this fact being used to define a standard of radioactive intensity. The amount of radon in equilibrium with one gram of radium is called the "curie," all radioactive intensities being referred to this or its sub-unit, the "millicurie." The volume of the curie can be calculated by measuring the number of a-particles given off per second by a gram of radium and determining the transformation constant of radon (Avogadro's number being assumed known) and by this means is found to be 0.59 cub. mm. This is in excellent agreement with the directly measured volume 0.60 cub. mm., as the smallness of the quantity makes a direct determination liable to error. Since the atomic weight of radium is 226, that of radon should be 222, if only one a-particle is emitted in the transformation from one to the other. This is the value found by direct density determination. affording thus strong evidence for the postulated mechanism of a-particle emission.

Analysis of the active deposits

A body exposed for a short time to radon coats with an active deposit which emits α -, β -, and γ -radiation and exhibits a regular decay. If the body is exposed to the radon for several days a residual a-activity results, increasing slowly over the first year and then decaying with a period of about 20 years. A great deal of experimental skill has been devoted to the problem of analysing the highly complex changes which occur. The analysis of the decay curves for both short exposure and long exposure to radon help to explain the nature of the transformations which take place. The conclusions so reached have been confirmed by making use of an ingenious method of separation by recoil. When a radon atom disintegrates, an a-particle is ejected and the atom which is left behind, called radium A (Ra A), recoils backwards. When radium A itself disintegrates a recoil atom, radium B, moves backwards, and so on for the whole of the series. Consider the recoil of the radium B atom which takes place when radium A emits its a-particle. The radium A disintegration results in the emission of an α-particle of mass 4 and velocity 1.82 × 109 cm./sec. As

the mass of the residual radium B atom is 214, its recoil velocity, from the conservation of momentum, is 3.4×10^7 cm./sec. When the radium B disintegrates to radium C, only a β -particle is ejected, and owing to the relatively small β -particle momentum, the resulting radium C recoil velocity is only 5×10^4 cm./sec., *i.e.* one seven-hundredth that of radium B.

The recoil of the residual atom can be used as a direct method for separating the disintegration products. These are always positively charged and if a negatively charged plate is placed close to the active matter, the recoil atoms can be captured and collected. They can then be removed and their individual decay periods determined, the values so found being in agreement with those deduced from the complex decay curves. By carrying out observations upon the deflections of the recoil atoms in magnetic and electric fields, values for the velocities and for E/M can be found. These results are in agreement with theoretical prediction. The recoil method, and the analysis of decay curves taken separately for the α , β , and γ activities of the active deposits, have enabled investigators to track out the whole series of complex transformations which take place.

The radioactive series

Although radium is *not* the parent of a series, the successive transformations following it will serve as an indication of the common type. The beginning of the sequence of changes following radium is shown below, the upper arrow indicating the radiation emitted in going to the succeeding atom in the series. Below each atom the period is shown

The equilibrium decay curve of Fig. 13.2 is based upon the above periods for radium A, radium B, and radium C.

When all the radioactive transformations are linked up into their respective series three main lines appear, beginning with uranium, thorium, and protactinium (which is the parent of actinium). The three series are independent, and there is some indication that the actinium series has a uranium isotope of mass 235 as its ultimate parent. Each series continues until a stable product is reached, but each also exhibits an occasional branching. The ejection of an α -particle leads to a reduction by 4 in the atomic weight of the remaining atom and by 2 in the nuclear charge. When a β -particle is emitted the atomic weight remains unchanged whilst the nuclear charge and atomic number increase by 1. The three series given by Rutherford are shown in Tables I, II, III. The names of some

THE RADIOACTIVE SERIES

TABLE I

Element		Atomic weight	Atomic number	Period	Sequence
Uranium 1	U 1	238	92	4.5×109 years	
Uranium X_1	$U X_1$	(234)	90	24.5 days	α
Uranium X ₂	U X ₃	(234)	91	1·14 min.	β
Uranium Z	U Z	(234)	91	6.7 hours	β
Uranium 11	UП	(234)	92	~106 years	
Ionium	Io	(230)	90	∼10⁵ years	α
Radium	Ra	(226)	88	1,600 years	α
Radon	Rn	(222)	86	3.8 days	
Radium A	Ra A	(218)	84	3.0 mins.	
Radium B	Ra B	(214)	82	26·8 mins.	α
Radium C	Ra C	(214)	83	19.7 mins.	β
Radium C'	Ra C'	(214)	84	1.5×10^{-4} sec.	β α
Radium C"	Ra C"	(210)	81	1·32 mins.	α ↓
Radium D	Ra D	(210)	82	~25 years	
Radium E	Ra E	(210)	83	5.0 days	β
Radium F	Ra F	(210)	84	136-3 days	Jβ
Radium G	Ra G	206	82	Stable	α

TABLE II

Elemen	ıt	Atomic weight	Atomic number	Period	Sequence
Protactinium Actinium	Pa Ac	(231) (227)	91 89	1.25×10^4 years 13.4 years	α
Radioactinium Actinium X Actinon Actinium A	Ac X An Ac A	(227) (223) (219) (215)	90 88 86 84	18-9 days 11-2 days 3-92 sec. 2-0×10 ⁻³ sec.	β α α α
Actinium B Actinium C Actinium C' Actinium C'' Actinium D	Ac C Ac C' Ac C' Ac C'	(211) (211) (211) (207) (207)	82 83 84 81 82	$36 \cdot 0$ mins. $2 \cdot 16$ mins. 5×10^{-3} sec. $4 \cdot 76$ mins. Stable	$ \begin{array}{c} \beta \\ \alpha \\ \beta \end{array} \qquad \begin{array}{c} \beta \\ \alpha \end{array} $

TABLE III

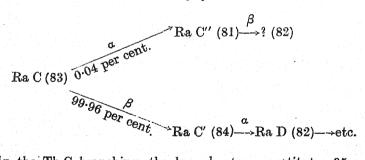
Element		Atomic weight	Atomic number	Period	Sequence
Thorium	Th	232	90	1.65×1010 years	
Mesothorium 1	Ms Th 1	(228)	88	6·7 years	α
Mesothorium 2	Ms Th 2	(228)	89	6·13 hours	β
Radiothorium	Ra Th	(228)	90	1.90 years	β
Thorium X	Th X	(224)	88	3.64 days	α
Thoron	Tn	(220)	86	54·5 secs.	∝
Thorium A	Th A	(216)	84	0·145 sec.	α
Thorium B	Th B	(212)	82	10.6 hours	α΄
Thorium C	Th C	(212)	83	60·5 mins.	β
Thorium C'	Th C'	(212)	84	$2 imes 10^{-7} { m sec.}$	α β
Thorium C''	Th C"	(208)	81	3-2 mins.	
Thorium D	Th D	(208)	82	Stable	β

atoms, such as Ra C' and Ra C', involve primes since they were only discovered after Ra D, Ra E, etc., had long been known and named.

The atomic weights in the brackets are deduced from the transformations only. Those of U1, Ra, RaG, Th, ThD, and AcD have been determined chemically, but are only given here as approximate whole numbers. The sequences described as β -ray transitions are usually accompanied by γ -radiation, a point which will be discussed later. It is clear from Table I, that radium is a member of the uranium series. Polonium, the early discovery of Mme. Curie, is radium F.

In the majority of cases an atom disintegrates in one way only but there are four exceptions which show branch products, the branching atom setting up its own disintegrating series. These can be seen in the series tables. The disintegration of Ra C is typical, the branching chain being as follows. It will be seen that the side branch to Ra C' constitutes only 0.04 per cent. of the total disintegration of Ra C so that the main chain is hardly affected at all. In three of the cases of branching the branch atom is produced only as a small fraction of the main disintegration.

Branching of Ra C

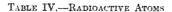


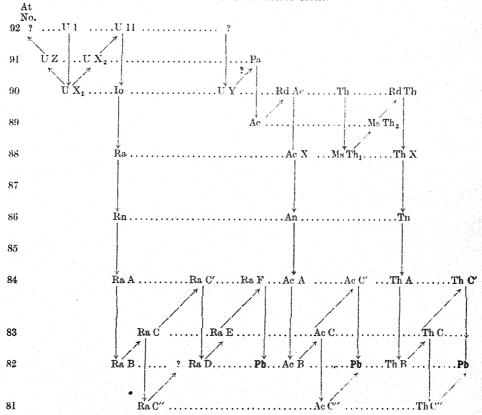
In the Th C branching, the branch atom constitutes 35 per cent. of the total disintegration.

Radioactive isotopes

The chemical properties of an atom are determined by its atomic number which fixes its place in the Periodic Table of the elements. Reference to the tables of the radioactive

series shows that frequent cases occur wherein atoms of different atomic weights have the same atomic number and thus occupy the same position in the Periodic Table. For example, U X_1 , U Y, Io, Rd Ac, Th, Rd Th all have different atomic weights and yet the same nuclear charge, 90. They constitute radioactive isotopes and, in fact, isotopes were first discovered by this means. In Table IV the radioactive atoms





are arranged according to atomic number and from this it is seen that isotopes are very frequent and also very numerous in some cases. All atoms which are isotopic are joined by a dotted line. Particular attention may be drawn to the elements with the atomic numbers 83, 82, and 81, for in the

Periodic Table the non-radioactive atoms of bismuth, lead, and thallium possess these atomic numbers and are therefore also isotopic with those radioactive elements of corresponding nuclear charges. This fact has proved to be of great value chemically, for if one of the radioactive isotopes of bismuth, say, is mixed with the latter, it behaves as an indicator and the bismuth can be traced by means of the associated radioisotope, as both behave identically to chemical and physical reactions. By this means the absorption of minute amounts of metal by materials, or by the body, etc., can be studied and a whole new field has been opened by the use of the "radioactive indicators."

The study of the chemical properties of the radio elements has led to the Displacement Law which states that when an a-particle is emitted the element shifts back its position in the Periodic Table by two places, whereas when a β -particle is ejected it goes forward one place. It is obvious from this rule that isotopes must result, as more than one type of atom can occupy a given place. The end product of each of the three series has atomic number 82, which is that of lead. Three stable lead isotopes of atomic weights 206, 207, and 208 arise from uranium, actinium, and thorium respectively. These are called uranium lead, actinium lead, and thorium lead, and are found separately in the minerals containing their parent atoms. This has been proved in a number of different ways. Direct chemical analysis of thorium and uranium minerals showed that the leads associated with these had the respective atomic weights 208 and 206. Ordinary lead is a mixture containing the three leads discussed above, as has been proved by the mass spectrograph and by hyperfine structure observations. elegant proof of the correctness of the radioactive series has been made by examining with high resolving power the spectra of the leads obtained from different parents. Although chemically identical (excluding atomic weight) the different leads exhibit a slight isotopic displacement in their line spectra and can be individually identified. The chemical determination of the atomic weights of the respective leads has completely vindicated the whole theory of the radioactive series, since the atomic weights of the leads can only be deduced after the analysis of the whole series has been completed for each.

Some properties of radium

Since radium has proved to be practically of great importance because of the high intensity of the radiations emitted, some of its properties will be considered here. Radium is present to the extent of 3.4×10^{-7} gram in each gram of mineral uranium. Its extraction therefore presents great difficulties. It is invariably found in uranium ores, being a member of the uranium series, and is separated by successive crystallisation of the bromide. Like any other atom, it has a characteristic line spectrum. Pure metallic radium melts at about 700° C. and is chemically analogous with barium. Since the first dis-

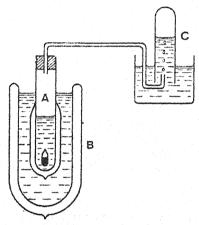


Fig. 13.3

integration produces radon, which is usually partially occluded in the radium compound, a mass of radium gives off radon in a gaseous form. Radium continuously maintains itself at a temperature a few degrees higher than that of the surrounding air, this heating effect remaining unaltered at very low temperatures. .

The first determination of the rate of emission of heat was made by Curie and Dewar with the apparatus shown in Fig. 13.3. The Dewar flask B contains liquid air and in it is immersed a second smaller Dewar flask A also containing The radium is enclosed in a small tube which warms up and communicates its heat to the thermally isolated

liquid air in the inner flask A. No heat can be transferred to the flask from outside, so that all the evaporation of liquid air in A is entirely due to the heat given out by the radium. The air is collected and measured, and from the latent heat of evaporation, the amount of heat generated can be calculated. It is found that 1 gram of radium in equilibrium with its products emits 140 grm. cals. per hour. It can be shown also that each atom of uranium in finally transforming itself to lead emits 7.1×10^{-5} erg, which corresponds to the energy acquired by an electron in falling through a potential difference of 45,000,000 volts. The heat produced by radium (or any radioactive body) is due both to the absorption of radiations and the energy of the recoil atoms. Most of the energy arises from α-ray absorption. The enormous quantity of energy radiated in radioactive transformations can only have its origin in nuclear sources. Indeed the high energies associated with the disintegration process first pointed to the nucleus as the origin of the effect.

REFERENCES

[&]quot;Radiations from Radioactive Substances." E. Rutherford, J. Chadwick and C. D. Ellis. (1930.)

"Radioactive Substances and their Radiations." E. Rutherford. (1913.)

CHAPTER 14

THE a-PARTICLE

The radioactive radiations can be distinguished by their different penetrating powers and by their different responses to the effect of strong magnetic and electric fields. The aradiations are completely absorbed by relatively thin metallic They can be deflected by strong magnetic and electric fields, which proves that they must be rapidly moving charged particles. It will be shown later that they consist of streams of high-speed helium atoms which have lost two electrons. The β -radiations are relatively much more penetrating than the a-radiations and are also much more easily deflected by magnetic and electric fields. Evidence will be given later showing that they are equivalent to cathode ray electrons of high speed. The y-radiations are extremely penetrating and cannot be deviated by magnetic or electric fields. They are light waves of extremely short wavelength, similar to X-rays but with a much shorter wavelength, and consequently more penetrating.

The ionising power of the respective radiations is associated with the degree of penetration possible, the more penetrating the radiations the less are they able to ionise. It is possible to separate the three types of radiation to some extent by means of suitable absorbing screens, but by this method only the γ -rays can be obtained quite free from admixture with the others. A complete separation is achieved by application of a strong magnetic field which deviates the positively charged α -particles to one side and the negatively charged β -particles to the opposite side, whilst the γ -rays pass on undeflected. The α -particles constitute the most energetic part of the radioactive radiations. This chapter will be devoted to the consideration of their properties.

Determination of E/M for the α-particle

The determination of the ratio of the charge E to the mass M was an important step in elucidating the nature of the α -particle. The principle of the method first used by Rutherford and Robinson is identical with that employed in making the determination of e/m for cathode rays. A beam of α -particles is deflected by magnetic and electric fields and from the displacements produced, E/M can be calculated. Absorption experiments show that the α -radiation given off by a mass of radium in equilibrium with its products is composite. If, however, a wire is exposed for some hours to radon, an active deposit of Ra A, Ra B, and Ra C is obtained, the decay periods of which are such that after a short time α -particles

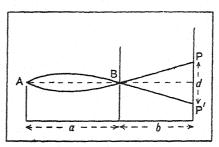


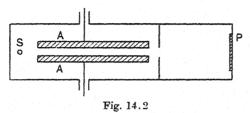
Fig. 14.1

are emitted from the wire by Ra C only. This therefore constitutes a homogeneous α -ray source.

The apparatus used for the magnetic deflection experiment is shown in Fig. 14.1. The source of a-particles, A, is set parallel to the slit B which is distant a cms. from it. At a distance b cms. from B is a photographic plate. The apparatus is evacuated and placed in a strong uniform magnetic field with the lines of force parallel to the slit, i.e. perpendicular to the plane of the paper. The field can be reversed after an interval. The a-particles, with velocity v, are deviated by the field and thus reach the points P and P' respectively before and after reversing the field. The calculation for relating the deflection with the field strength is identical with that carried out for the cathode rays giving 1/r = HE/Mv, H being the field strength and r the radius of the curve through which the particles are deflected. From the geometry of the apparatus

it can be shown that, for small deflections, 2rd = a(a+b), hence the measurement of the magnetic deflection gives the value of E/Mv.

In order to be able to calculate E/M an electric deflection experiment must now be carried out. When the α -particles are allowed to pass between parallel plates maintained at a different potential, the deflection produced is small and difficult to measure unless the deflecting path is long. For long paths the radiation given off by a coated wire is insufficient to produce a photographic effect. The wire must therefore be replaced by a tube containing a large quantity of radon, the tube walls being very thin to reduce absorption effects. The experimental arrangement used by Rutherford and Robinson is shown in Fig. 14.2. The α -rays from the source S pass between the deflecting plates AA which are 35 cms. long and 4 mm. apart. After passing through a mica slit one-sixth of a



millimetre wide the rays travel 50 cms, to P, a photographic plate. The whole apparatus is evacuated and a potential of 2,000 volts applied between the plates AA. The treatment is again similar to that in calculating the cathode ray electrostatic deflection. The rays have a parabolic path between the plates, the deflection being proportional to $1/v^2$. By reversing the voltage on AA, symmetrical deflections are obtained, leading to a high degree of accuracy in measurement. Since the radon source radiates a-particles from radon, Ra A and Ra C, three deflection bands appear upon the photographic plate. These particles are separated from each other since they have different velocities and are thus deflected differently. This introduces no complication in the experiment as long as the field strength is sufficient to resolve the individual particles. A composite source can be used in the magnetic deflection experiment too, providing the resolution is great enough.

The electrostatic deflection experiment enables the value of

E/Mv2 to be calculated and as the magnetic experiment gives E/Mv, both E/M and v are obtained by combining these. The value of E/M found for all a-particles, irrespective of the source, is 4,820 e.m.u. This is very close to the value for E/M calculated for doubly ionised helium, the latter being 4,826 e.m.u. Identity was therefore suspected and this was proved in a decisive manner by Rutherford and Royds, using the following method. The a-particles given off by a large quantity of radon enclosed in a thin-walled tube were allowed to gather in a highly evacuated discharge tube. After two days a spectroscopic examination of the tube showed that helium had appeared in it. The helium lines gradually became more and more intense with increasing time. It had long been known, however, that radioactive ores and radium itself occluded helium. Control experiments were therefore carried out to prove that no occluded helium was able to pass through the walls of the tube containing the radon. It is thus quite evident that the a-particles themselves, which do pass through the tube walls, give rise to the observed helium spectrum. In the process of absorption by the walls of the discharge tube the a-particles capture electrons and become normal helium gas atoms.

The helium found occluded in radioactive ores arises from the absorbed α -particles. A further proof of the identity with helium is obtained by firing the α -particles into a sheet of lead which absorbs them. Electrons are captured by the Particles in the absorption process and they become normal helium atoms. If the lead is melted, this helium is evolved and can be detected by spectroscopic methods.

The electrostatic and magnetic deflection experiments give also the velocity of the particles. These are found to be very high indeed, differing for different sources. For example, the velocity of the α-particles from Ra C is 1.922×109 cms./sec. which is one-sixteenth of the velocity of light. The penetrating power is the direct result of the high velocity, for the speed reduces the chances of losing energy in collision with an atom when passing through matter.

The detection of a single α -particle

The a-particles are extremely energetic, those from Ra C, for instance, having each a kinetic energy of 1.2×10^{-5} erg.

This corresponds to the energy acquired by an electron in falling through 7.66 million volts. The energy associated with each a-particle is relatively so great that, despite their atomic character, individual particles can be detected. Four methods have been employed for the detection of the particles as individuals, these being conveniently described as electrical, optical, photographic, and expansion chamber. The electrical method depends upon the fact that in an average case, as in the emission from Ra C', a single a-particle can produce as many as $2\cdot2\times10^5$ pairs of ions in its path through air. As each of the ions formed has a charge of $4\cdot8\times10^{-10}$ e.s.u. this corresponds to a total quantity of about 1×10^{-4} e.s.u. of electricity. Although this quantity can be detected it is small, and in order to magnify it Rutherford and Geiger made use of ionisation by collision in a very ingenious manner. The

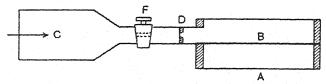
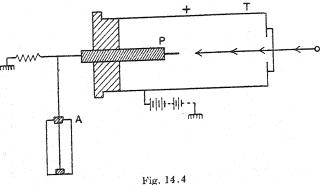


Fig. 14.3

device employed for the detection of the individual particles is shown in Fig. 14.3. The active material in C sends out a stream of a-particles and as the vessel is evacuated these reach the circular hole D, which has a diameter of 1.5 mm. The number of a-particles reaching D can be cut down by the stopcock F to about three per minute. D is covered with a very thin sheet of mica, which although it has only a small stopping power, enables the vessel C to be evacuated whilst the pressure in the vessel AB, the detecting chamber, is maintained at a few centimetres of mercury. The detecting vessel consists of a metal cylinder A, along the axis of which passes an insulated wire B. The gas pressure is adjusted so that if a voltage is applied between the wire and the cylinder the ions produced by the entry of an a-particle are enormously increased by ionisation by collision and the momentary current produced can quite easily be detected. With this apparatus the total number of particles given off by a known quantity of radium

in 1 second can be counted. If a string galvanometer is used as the current detector and photographic registration employed, as many as 1,000 particles per minute can be counted, if a suitable leak is used to permit the string to come back rapidly to zero.

The α -particle counter just described has been improved by Geiger and made so sensitive that even the very much weaker ionisation produced by a single β -particle can be detected. The modified Geiger counter is shown in Fig. 14.4. The rod P has a fine, pointed end and projects into the brass tube T, being carefully insulated from it. The particle enters through a window which is covered with a thin mica sheet so that the pressure within the tube T can be adjusted. A voltage is applied between T and P and maintained very near to the



discharge point. The entry of a particle causes local ionisation near the point and this has a trigger action effect, causing a short discharge which is recorded by a self-registering string galvanometer A. The actual magnitude of the discharge current seems to be independent of the initial ionisation, and owing to this, the instrument will detect β -particles as well as a-particles. The mechanism of the counter is quite different from that of the one previously described and cannot be used if a-particles are to be examined in the presence of β -particles. It is possible to modify the counter by using a small sphere instead of a fine point, so that only a-particles give a response.

The optical or scintillation method of observing individual particles, due to Crookes, has proved to be a sensitive method of great value. When a screen is coated with the phos-

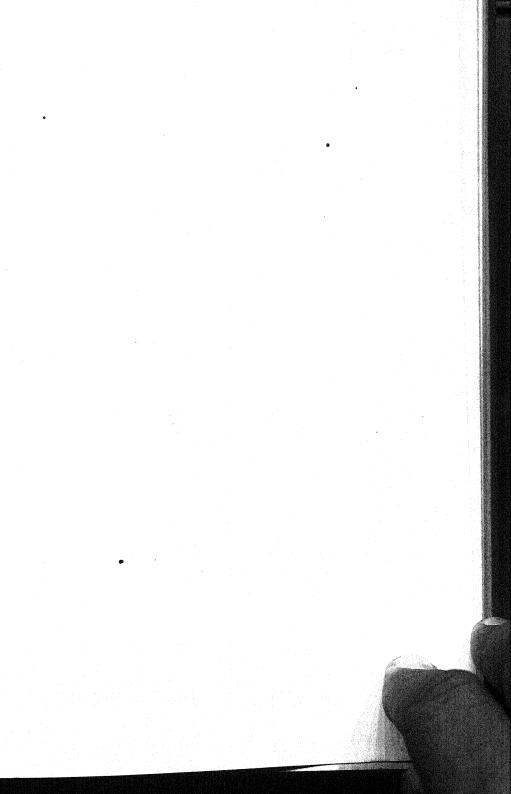
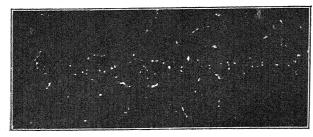
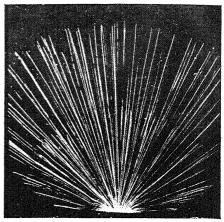


PLATE IV



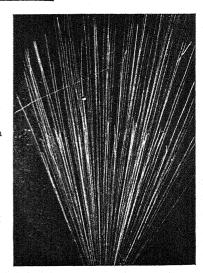
IVA. Ion track produced by X-ray beam (after Wilson).

IVB. a-particle tracks of two distinct ranges (after Rutherford, Chadwick and Ellis)



IVc. a-ray tracks, one exhibiting nuclear deflection (after Wilson).

IVd. Nuclear disintegration by a-particle with proton emission (after Blackett).



phorescent crystals of zinc sulphide and exposed to a source of α -radiation it becomes luminous, and when viewed with a lens this luminosity is seen to consist of a number of scintillating points. Each flash is due to the impact of a single α -particle upon a crystal. A delicate means for counting the individual particles is thus available.

The third method, the photographic, will be very briefly dealt with for its applications are very limited. It is possible to observe the photographic tracks of the a-particles if they are allowed to fall tangentially upon a photographic plate, or by locally infecting the plate itself with radioactive material. In the latter case the particles passing through the gelatine in a direction parallel to the plate leave tracks which can be developed up.

The most powerful method of studying the individual particles is by the use of the Wilson expansion chamber. An example of Wilson chamber photographs is shown in Plate IVB and it will be seen from this that the particles travel in straight lines and occur in groups, the range having very nearly a constant value in each group. The occasional sudden deviations from the straight-line track which can be seen near track ends are due in each case to the collision of an a-particle with the nucleus of a gas atom in its path (Plate IVc). Sometimes, when the conditions are suitably adjusted, short electron tracks can be seen emerging from the main a-particle track. The electrons responsible for these tracks have been called δ-rays. They are in fact the secondary electrons of highest energy produced as a result of the ionising activity of the a-particles. The use of the expansion chamber has yielded a number of a very important facts particularly concerning the disintegration of atomic nuclei.

The charge carried by an α-particle

The value of E/M for the a-particle indicates that the particle is to be identified with a doubly ionised helium atom. A final proof of this is forthcoming if the actual charge E can be separately determined, as this enables both E and M to be calculated and the identity to be completely established.

E can be most conveniently measured if the charge carried by a known number of particles is ascertained. There exist a number of methods for determining the number of a-particles given off in 1 second by a gram of radium. It can be measured by means of a counter, by the direct observation of scintillations, and by collecting and measuring the amount of helium given off in a known time (156 cub. mm. are given off in a year). As a mean of these it is found that 1 gram of radium emits 3.70×10^{10} a-particles per second.

Since the number of particles given off per second is known,

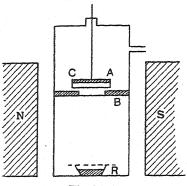


Fig. 14.5

E can be determined by measuring the total charge given up by a known number. The method of doing this is illustrated in Fig. 14.5. A known quantity of Ra C is contained in a shallow dish R which is covered with very thin aluminium foil to keep back recoil atoms. The α -particles fall upon the collecting plate CA which is connected to an electrometer. The collecting plate is covered with thin aluminium foil, and its area is defined by the diaphragm B. From the area of the diaphragm and the known emission by the source, the number of α -particles falling on the plate can be deduced, and so the charge per particle measured. The whole apparatus is maintained in a strong magnetic field (N, S are the pole pieces of a powerful magnet) the objects of which are twofold. Firstly, it deviates away the β -rays given off by the source preventing them from reaching the collector. Secondly, it

curves back on to the collector the secondary electrons liberated from the surface of the metal by the impact of the α -particles. Thus the only charge recorded on CA is that brought by the α -particles. The apparatus is evacuated in order to avoid complications which would be produced by collisions, etc. The value found for E is 9.3×10^{-10} e.s.u., which is twice the electronic charge. Thus within the limits of experimental error it is confirmed that the α -particle is a helium atom which has lost its two outer electrons. It is a high-speed helium nucleus, since the normal helium atom has only two outer electrons.

The absorption of α-particles by gases

One method of studying the absorption of α -particles in gases is to measure the length of the track in an expansion

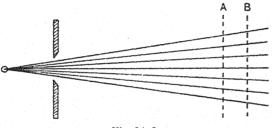
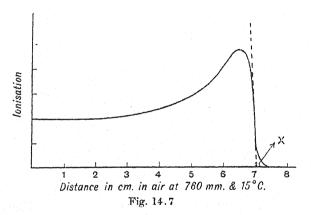


Fig. 14.6

chamber filled with various gases at different pressures. A more precise method is due to W. H. Bragg, the details of the method being outlined in Fig. 14.6. A narrow pencil of a-particles is formed by allowing the radiations from the source to emerge from a small aperture in a lead box. Two metallic gauzes A and B, parallel and close to each other, form an ionisation chamber. The saturation current between these gauzes is proportional to the intensity of the ionisation produced between them by the a-particles. The source can be moved towards and away from the gauzes and by this means the amount of ionisation produced at every point of the a-particle path can be measured. Alternatively the source and ionisation chamber can be kept fixed whilst the gas pressure is varied.

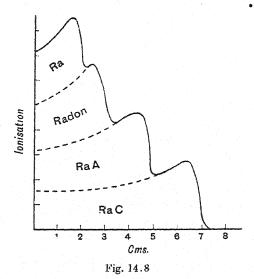
With a homogeneous source of α -radiation, such as that given off by Ra C', Bragg found that the curve relating the ionisation with the distance from the source is as illustrated in Fig. 14.7. This curve shows that the ionisation along the path of a particle climbs steadily to a maximum and then falls with great rapidity to zero, making a slight ankle in the curve just before the zero line is reached. If radium in equilibrium with its products is examined, the α -radiation contains α -particles emitted from Ra, Ra A, radon, and Ra C, and as these have different ranges, the resultant ionisation is the sum of the ionising effect of each of the different types of particle. Each different group produces a curve similar in



shape to that of Fig. 14.7, with, however, the scale differing in each, since the zero (end of the range) is reached at different points. The curves of ionisation are therefore additive, leading to the resultant thick curve in Fig. 14.8. The dotted continuations show how the ionisation curves of the individual groups add up to give the resultant for the complex radiation.

The ionisation curves may be used to determine the ranges of the particles, since they indicate the point at which the particles lose their power to ionise. The tail-end ankle of the curve in Fig. 14.7 (the straggle effect) arises from the fact that the collisions of the α -particles with atoms obey the laws of probability, with the result that individual ranges vary slightly from the average. The range is taken to be the extrapolated

value found by producing the almost straight falling portion of the curve to the x axis. The distance to which the straggle tail stretches depends upon the sensitivity of the measuring device used; the higher the sensitivity the more easily can the few particles with excess range be detected. By measuring the amount of ionisation over the whole range, that is, by integrating the ionisation curve, it can be shown that the total

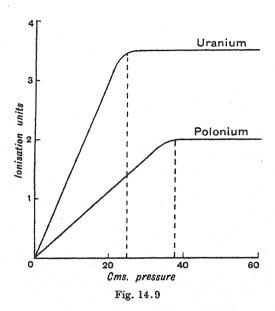


number of ions produced in air at 760 mm. pressure and 15° C. by the absorption of a single a-particle from, say, Ra C' (range 6.97 cms.) is 2.2×10^5 .

The ranges of α -particles

The ranges of the particles can be deduced from the ionisation curves in the manner described, but this method is not well suited to weak sources, for which a simple method has been devised by Geiger and Nuttall. With this a very complete study of α -particle ranges has been made. The active material is placed on a rod at the centre of a glass bulb which is silvered on the inside, and a high voltage is applied between the bulb and the metallic holder of the active material. The saturation current for different gas pressures is measured,

giving ionisation curves similar to that shown in Fig. 14.9. Starting with low pressures, the current rises linearly with increasing gas pressure until a point is reached where all the a-particles are absorbed and thus achieve complete ionisation. Further increase of pressure now is without effect on the measured current. At the sharp turn-over point in the curve, the a-particles just reach the silvered wall of the bulb, so that the bulb radius is the range. The range for a particular gas pressure has thus been measured, and since the range in a gas



is inversely proportional to the pressure, the range at atmospheric pressure is simply deduced. Geiger has improved the method and rendered it capable of giving ranges with an accuracy of one-hundredth of a millimetre.

The ranges of the different a-particles in air may also be interpreted in relation to the stopping power exerted by the air. A given range means that a given number of centimetres of air are required for absorption of the particles. If a sheet of any material is interposed in the path of the particles it exerts a retardation effect and reduces the range. The stopping power of the sheet is defined as the equivalent centimetres

of air path by which the normal range has been reduced. The stopping power of a sheet depends upon its thickness, atomic weight, and density, and to some extent upon the actual initial velocity of the α -particles which are being absorbed. To a first approximation it is proportional to the square root of the atomic weight (Bragg and Kleeman's law).

When the a-particles pass through matter the majority travel in straight lines losing energy by removing electrons from atoms lying along the whole length of the path. On rare occasions a collision with the relatively small nucleus of an atom takes place, in which case the a-particle suffers a large angle deflection sometimes exceeding 90°. Since energy is used up in producing ionisation along the whole length of the path, it is obvious that the velocity of the particle must continually diminish after it has left its parent atom. (This is not very marked in a vacuum where there are relatively few collisions.) A relationship exists between the velocity retained after emerging from an absorbing sheet of material and the residual range left. This, known as Geiger's law for the reduction in velocity, can be expressed as $V^3=aR$ where a is a constant, V the velocity of the particle after leaving the absorbing material, and R the range in air after passing through the absorber. Geiger derived this law by measuring the velocities of the emerging particles by the magnetic deflection method. It does not hold quite accurately for the longer range particles.

The fine structure of α -particle ranges and long range α -particles

It was for a long time considered that the a-particles emitted by a single radioactive body all had the same range, excluding the straggle effect due to statistical fluctuations of the collisions in the absorbing medium. However, an indication that this was not so was first found by Rutherford in Th C'. For every 10⁶ a-particles emitted with the normal range of 8.6 cms. there are 35 particles with a 9.7-cms. and 180 with a 11.6-cms. range. Although these long-range particles are few in number they come from the same nucleus as the rest. They are clearly more energetic, and it can be shown that they arise from excited levels within the nucleus itself.

When a main group of α -particles is more closely examined, it is sometimes found to exhibit a fine structure, that is to say, to consist of a number of discrete individual groups of particles whose ranges are very nearly the same but are definitely distinct. Rosenblum first observed this in the α -radiation given out by the disintegration of Th C into Th C". The α -particle beam was deflected by a very large magnet weighing 120 tons and producing a uniform field of 36,000 gauss over a circular gap 25 cms. in diameter and 3 cms. wide. Photographic registration was used. Later Rutherford and collaborators used a fixed radius of curvature and varied the magnetic field, bringing particles with different velocities successively into an ionisation chamber. By these means the fine structure of the α -particles from several bodies was observed.

The strength of the magnetic field is the main factor enabling this fine structure to be resolved. With the very large fields employed it is possible to separate particles whose velocities differ by only 0.02 per cent. The fine structure can also be correlated with excitation energy levels within the nucleus, but the levels in this case are those of the product nucleus, not the nucleus from which the particles are emitted.

α -particle range and the transformation constant

A relationship of practical application has been found to exist between the values of the transformation constants and the a-particle ranges for different radioactive materials. In general when the transformation constant is large (and the period therefore small) the a-particle range is also large. This relationship, known as the Geiger-Nuttall rule, is best shown by plotting the logarithm of the range, $\log R$, against the logarithm of the transformation constant, $\log \lambda$. The plot for the members of the three radioactive series shows that three nearly parallel straight lines are formed, one for each of the series. The graphs are not exactly straight lines, but to a first approximation they can be fitted into equations of the form $\log \lambda = a + b \cdot \log R$, the constant b being the same for the three series whilst a has three different values.

It has been found possible to derive a formula in justification of the Geiger-Nuttall rule. Theoretical wave mechanics methods

show that the energy of an a-particle, E, is related to the transformation constant λ according to the law $\log \lambda = a + b$. E. This corresponds to the empirical Geiger-Nuttall law since the a-particle range depends upon the energy. The theory shows, too, that b is not quite constant for all different materials but decreases when E increases. This accounts in part for the deviations from the straight line which are found when the respective logarithms of transformation constant and range are plotted against each other. The rule has proved to be extremely valuable for fixing roughly the transformation constants of very short or very long lived products. All that is required is that the a-particle range be measured and from this an approximate value of λ is easily calculated.

The scattering of α -particles by matter

Fundamental discoveries concerning our knowledge of atomic structure have developed from Rutherford's investigation of the scattering of a-particles by matter. If a sharply defined pencil of a-rays is allowed to fall upon a photographic plate, in vacuo, the photographic image which is formed has clean, sharp edges. If a thin screen with a stopping power of about a centimetre is interposed in the beam, the image broadens out and becomes diffuse, due to scattering of the particles by the atoms of the absorbing screen. The majority of the a-particles are only scattered through small angles of a few degrees but some behave quite differently. The latter are deflected through very large angles, even up to 150°. The number of such particles, which are almost "reflected" rather than deflected, increases with the thickness of the absorbing screen until a certain limiting thickness is reached, after which they remain constant. The number also increases rapidly with the atomic weight of the scattering screen, as many as I in 8,000 being turned through more than a right angle by a thick sheet of platinum foil.

From measurements of the number of particles scattered through small angles, it is certain that the number scattered through very large angles is abnormally high if the law of chance is obeyed. It must be concluded that the scattering mechanism for large and small angle scattering is not the same. Rutherford showed that the only possible explanation of the large angle

scattering is that it is due to a single encounter between the a-particle and an atom possessing a very intense central electrical field. It is impossible to explain the relatively high frequency of large angles by assuming that successive small angle scattering of one particle has taken place. In order to meet the requirement of such an intense central field, Rutherford proposed an atomic structure which has now become classical. According to this view, already discussed, the atom consists of a very small massive positively charged nucleus, in which most of the atomic mass is concentrated. Around this and outside of it is an equivalent negative charge distributed over a sphere. This we now identify with the outer rotating electrons of the atom. The whole atom is therefore neutral. but as the nucleus is considered to be very small compared with the radius of the sphere containing the negative charge, there is effectively a positive point charge at the centre. The physical radius of the atom is the radius of the sphere containing the negative charge, which in effect is equivalent to the radius of the outer electronic orbit. Rutherford proved that for scatter deflections greater than 1° the outer negative charge due to the electrons can be neglected entirely, the nucleus being the only agent responsible for the scattering. The method of calculating the nuclear scatter effect, which is due to Rutherford, will now be considered.

For simplicity it is assumed that an oncoming α -particle is effectively a positive point charge associated with a given mass and velocity. Owing to its great speed it can penetrate the outer electron cloud and approach closely to the positively charged nucleus of the atom. An electric force of repulsion, obeying the inverse square law is set up between the atomic nucleus and the α -particle, since they have like charges. If the atomic nucleus is relatively heavy and assumed to remain at rest, the α -particle will describe a hyperbola about the nucleus as external focus, according to the following treatment.

Consider an α -particle moving along PO (Fig 14.10), approaching the heavy nucleus which is stationary at S. It is then deflected into the path OP'. Let p=SN be the perpendicular distance from the nucleus S to the original direction of the oncoming particle. Let E, M, V, be the charge, mass, and initial velocity of the α -particle. The charge upon the nucleus of the

atom at S is Ze where Z is the atomic number of the atom. The velocity of the α -particle varies along its path. Let it have the value v at the point A, which is the closest distance it approaches the nucleus at S. The particle is moving perpendicularly to SO when it is at A, hence from the conservation of angular momentum p.V = SA.v. From the conservation of energy the initial kinetic energy is equal to the sum of the potential and kinetic energies at the point A, so that

$$MV^2/2 = Mv^2/2 + ZeE/SA$$

 $v^2 = V^2(1-b/SA)$ where $b = 2ZeE/MV^2$.

whence

With a head-on collision a particle is instantaneously brought

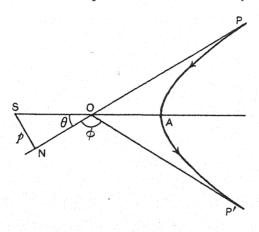


Fig. 14.10

to rest at the point where its initial kinetic energy equals the potential energy acquired. If this closest approach equals x we get $MV^2/2 = ZeE/x$, therefore $x=2ZeE/MV^2$ which is identical with the value for b used above. As the eccentricity of the hyperbola is $\sec \theta$, then from the geometry

$$\begin{array}{ll} \mathrm{SA}\!=\!\!\mathrm{SO}\!+\!\mathrm{OA}\!=\!p\,\,\mathrm{cosec}\,\,\theta\,\,(1\!+\!\cos\,\theta) \\ =\!p\,\,\mathrm{cot}\,\,(\theta/2) \\ \mathrm{Thus} \qquad p^2\!=\!\!\mathrm{SA}(\mathrm{SA}\!-\!b)\!=\!\{p\,\,\mathrm{cot}\,\,(\theta/2)\}\{p\,\,\mathrm{cot}\,\,(\theta/2)\!-\!b\} \\ \mathrm{therefore} \qquad b\!=\!2p\,\,\mathrm{cot}\,\,\theta. \end{array}$$

The angle ϕ through which the α -particle is deflected is $\pi-2\theta$ herefore $\cot (\phi/2)=2p/b$.

It remains now to calculate what will be the probability of the deflection of a particle through the angle ϕ . Suppose the incident pencil of α -particles falls normally upon a thin screen containing n atoms per c.c. The thickness, t, of the screen is small enough not to have any appreciable effect upon the velocities of the α -particles passing through. The probability, q, of an α -particle passing within a distance p of a nucleus is given by $q = \pi p^2 nt$. Such a particle will be deflected through an angle greater than ϕ . As q is a probability, it represents the fraction of the total which is deflected through an angle greater than ϕ , and by substituting for p we get

$$q = \frac{1}{4}\pi ntb^2 \cot^2(\phi/2).$$

Since the probability of being deflected between ϕ and $\phi+d\phi$ is the same as that of passing the nucleus within a region between p and p+dp

then

$$dq = 2\pi pntdp$$

$$= \frac{1}{4}\pi ntb^2 \cot (\phi/2) \csc^2 (\phi/2) . d\phi.$$

If the total number of particles falling upon the absorbing screen is Q, the number deflected between ϕ and $\phi+d\phi$ will be

$$N = \frac{1}{4}\pi Qntb^2 \cot (\phi/2) \csc^2 (\phi/2) \cdot d\phi$$
.

In carrying out the experiments the scattered particles are made to produce scintillations by falling upon a zinc blend screen which is always maintained perpendicular to the direction of viewing. By this means the number of particles, incident normally upon the screen, which have been deflected through the angle ϕ can be counted. Let M' particles be observed per unit of area at a distance r, then it follows that

$$N=2\pi r \sin \phi . r d\phi . M'$$
.

Account is thus taken of all the particles lying between the cones with half angles ϕ and $\phi+d\phi$. The two values obtained for N can now be equated giving

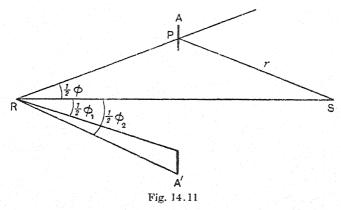
 $\frac{1}{4}\pi Qntb^2 \cot (\phi/2) \csc^2 (\phi/2) . d\phi = 2\pi r^2 \sin \phi . d\phi . M'$ therefore

$$\mathbf{M'} = \frac{\mathbf{Q}ntb^2 \cot (\phi/2) \cdot \operatorname{cosec}^2 (\phi/2)}{8r^2 \sin \phi}$$

which can be rewritten as

$$M' = \frac{Qntb^2}{16r^2} \csc^4 (\phi/2).$$

Since $b^2=4Z^2e^2E^2/M^2V^4$ it follows that the number of a-particles falling upon unit area of the scintillation screen is proportional to: (1) $\csc^4(\phi/2)$; (2) t; (3) the square of the nuclear charge Ze; (4) inversely as V^4 . Geiger and Marsden carried out an experimental test of the theory by measuring the distributions of scattered particles with scattering screens of different thicknesses and of different materials. They found that the scattering was proportional to $\csc^4(\phi/2)$, to t, and also to the square of the atomic weight. Since, to a first approximation, the nuclear charge is proportional to the atomic weight these experiments verified the formula deduced



by Rutherford and so proved that the assumptions made were justified.

The exact verification of the third point above, namely, that the scattering is proportional to Z^2 , was made by Chadwick, who at the same time was able to make an absolute determination of the nuclear charge of a number of atoms. In order to increase the number of observable particles scattered through large angles, Chadwick used a fairly wide annular ring of foil as the scattering screen. This subtends a wide cone at the source R and the scintillating screen S (see Fig. 14.11). The distances RA and SA are made equal, and it can be shown that the number of scattered particles falling upon unit area at S, perpendicular to RS, is

$$\begin{array}{l} \{Qntb^2/64r^2\}\{\log\,\tan\,\phi_2/4 - \log\,\tan\,\phi_1/4 \\ +\cot\,\phi_1/2\,\csc\,\phi_1/2 - \cot\,\phi_2/2\,\csc\,\phi_2/2\} \end{array}$$

The angles ϕ_1 and ϕ_2 have the meanings indicated in Fig. 14.11; r is the mean value of the distance between the scattering foil and the source, i.e. RP, and n, b, t have the same meanings as before. Q, being the total emission of particles per second from the source, can be determined by measuring the direct number of particles striking unit area of the fluorescent screen at S. Since the direct beam is about a thousand times more intense than the scattered beam, it can be cut down to a suitable known ratio by a rotating sector in order to give a number of scintillations which can be reasonably counted.

From the above expression for the number of particles falling upon unit area, it is possible to derive a value for b. All the quantities in the formula for b are known except Z_{ϵ} , the nuclear charge of the scattering atom, hence the value obtained for b enables an absolute determination of Ze to be made. This method is at present the only one available for making an absolute determination of nuclear charge and is therefore important. Chadwick measured Z for copper, silver, and platinum, obtaining the values 29.3, 46.3, and 77.4 respectively. These atoms occupy the 29th, 47th, and 78th positions in the Periodic Table, the atomic numbers being therefore 29, 47, and 78. Thus the values derived for Z by means of Rutherford's scatter formula agree with those given by the Periodic Table. From this it is clear that Rutherford's theory of atomic structure is upheld. It is further possible to show from the scattering measurements that the inverse square law . . of repulsion between the nucleus and a-particle must hold up to a distance of approach of the order of 10^{-12} cms. This proves that the nucleus is minute in comparison with the atomic diameter which is of the order of 10-8 cms.

Some general properties of α -particles

All the three radiations, α , β , and γ , are able to induce physical and chemical changes in a number of materials. Glasses, crystals, and many minerals are changed in colour when subjected to irradiation by any of these rays, some being very sensitive indeed. Glass and quartz become brittle and mica sheets discolour and even bend if intense a-radiation is employed. The a-particles are able to decompose water and

can produce a number of chemical changes. Related to the chemical activity are the well known physiological effects. These are largely due, however, to the β - and γ -rays, since α -particles are completely absorbed near the surface of the body. The penetrating γ -radiation is able to destroy body cells, attacking the malignant cells more easily than the normal. For this reason a considerable degree of success has been reached in the destruction of malignant tumours by the local application of tubes containing radium or radon.

The fact that a-rays are able to turn mica brown has produced an interesting natural phenomenon. Geological specimens of mica are occasionally found to exhibit small discoloured areas, circular in section. Since they show the optical phenomena of pleochroism when examined with polarised light, they are called "pleochroic haloes." It can be demonstrated that a small speck of radioactive impurity causes the halo. A centre dark patch is produced by the a-particles from radium, its radius corresponding to the range of these particular particles in mica. A second dark ring is formed by the a-particles given off by the radium A naturally present, whilst a third outer dark band forms, due to the longer range radium C'a-particles. The degree of darkening at any point in the range depends upon the amount of ionisation produced. We know from Bragg's experiments that this rises rapidly to a maximum near the end of the range, a fact which explains why distinct rings are formed and not continuous discs.

It can be shown that a mica halo can be produced with as 5×10^{-10} gms. of uranium as the central exciting material. Such a quantity of uranium emits only one aparticle every 10 hours. The observed haloes are due to the accumulated integrated effect over immense periods of time. Joly has compared the coloration intensities found in nature with those produced by a known very large number of aparticles given off by a strong source. He was able to prove that several hundred millions of years are needed for the production of natural haloes. This time estimate has proved to be of great value to geologists and geophysicists.

Haloes occur in many materials other than mica. Some samples of blue fluospar have been found exhibiting seven

ringed haloes, each ring arising from an a-particle group with a distinct velocity. Artificial haloes can be produced in glass or upon photographic plates and are similar to those occurring naturally.

Another striking property of a-rays is that they are able to excite luminous effects by impacting upon bodies. The scintillations already mentioned are but a particular application of this property. Quite a large number of minerals and salts glow under a-ray bombardment. This has been utilised in industry, luminous paints being manufactured by mixing zinc sulphide with a small quantity of a radioactive body. It can be shown that bombarded solids not only emit light, but also send out weak X-rays. A phenomenon related to the production of scintillations is the fact that a mass of radium viewed in the dark is seen to glow. If the radium is in air spectroscopic examination shows that the glow consists of the spectrum of nitrogen. It is only due to the optical excitation by the impact of the radiations on the gas surrounding the radium.

REFERENCE

[&]quot;Radiations from Radioactive Substances." E. Rutherford, J. Chadwick and C. D. Ellis. (1930.)

CHAPTER 15

THE β-RAYS

Introduction

Although the identity of the β -rays given off by radioactive materials was established at an early stage in the development of the subject, their complexity made an analytical study very difficult. The Curies were the first to show that some of the radioactive substances emit negative ions which are much more penetrating than the α-particles. These are the β-rays which Becquerel found easy to deflect in a magnetic field. This magnetic deflection proved that the rays consist of charged particles. In Becquerel's experiment the rays were allowed to emerge from a slit and after being deflected into circles by a magnetic field fell upon a photographic plate which they blackened. The plate showed, instead of a sharp image of the slit, a broad diffuse band. If the particles are assumed all to have the same mass and charge this broadening can be readily explained by assuming that the beam consists of a mixture of particles with different velocities. As a result of this the slower moving particles are more easily deflected than the faster (radius of curvature of deflection R = mv/He). hence a whole band is covered by the particles.

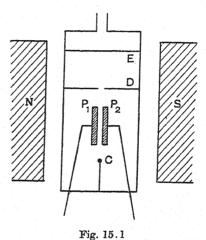
The nature of the charge carried by the β -ray particles was demonstrated in a simple manner by allowing the radiation to impinge upon a thick lead plate connected to an electroscope. The particles were absorbed by the plate and gave up their charges, which proved to be negative. In order to avoid any secondary effects produced by the total radiations given off by the radioactive material, the experiment was conducted in a high vacuum. The particles appeared indeed to behave like fast cathode-ray particles, that is, like high-speed electrons, and this identity was definitely established when it was shown that

249

the value of e/m for the β -particles was the same as that for the electrons.

Determination of e/m for the β -ray particles

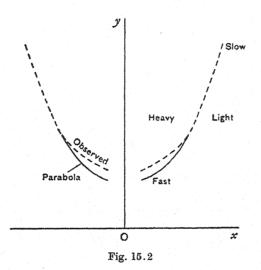
Approximate measurements for e/m for the β -particles were made by Becquerel, but it was Kaufmann who first obtained an exact value and proved the identity of the electron with the β -particle. Kaufmann's experimental arrangement is shown in Fig. 15.1. A small piece of radium is placed in a highly evacuated vessel at the point C and the radiations from this are allowed to pass between the parallel plates P_1 , P_2 . The



plates are 0·15 cm. apart and are maintained at a potential difference of 6,750 volts. The rays pass through a hole in D a diaphragm, and then strike a photographic plate at E. The γ -radiations are undeflected by the electric field and as the deflection produced on the α -rays is very small indeed with the field employed, these radiations result in an undeflected zero spot upon the photographic plate. The effect of the field upon the β -particles is to deflect them to the right. The whole apparatus is maintained in a uniform magnetic field NS which is so arranged as to deflect the β -particles in a direction at right angles to that produced by the electrostatic field, *i.e.* in a direction perpendicular to the plane of the paper. The

electrostatic deflection is proportional to Xe/mv^2 and the magnetic deflection to He/mv where X, H, e, m, v have the usual meanings. These two deflections are at right angles to each other, and if v varies continuously a parabolic curve will result on the photographic plate. By measuring the coordinates for any point on the curve both e/m and v can be calculated, exactly as in Thomson's determination of E/M for positive rays.

Kaufmann found that the β -particles given off by the source he employed had very high velocities, varying from



 2.83×10^{10} cms./sec. to 2.36×10^{10} cms./sec. These velocities are of the same order as the velocity of light. Exact measurement proved that the deflection curve was not a true parabola but of the shape shown in Fig. 15.2. For the purpose of exact measurement it is convenient to reverse the magnetic field in order to obtain the symmetrical pattern shown. The observed points are plotted upon the dotted line, the continuous line indicating the true parabola. The faster moving particles are those nearest the origin, being the least deflected by the fields, and the deviations from the parabola show that e/m is not constant but diminishes as v increases. This had been predicted by Lorentz from the following considérations.

A charge may be regarded as reaching out tubes of force into space, the tubes having energy associated with them. In order therefore to set a charge and its attendant tubes in motion work must be done upon it. The charge has, therefore, an apparent inertial mass, called the electromagnetic mass. The equivalent mass of the tubes can be calculated in a simple manner since the moving charge produces a known magnetic field and the electromagnetic energy per cubic centimetre in a field H is known to be $\mu H^2/8\pi$ where μ is the permeability. The calculation shows that the total electromagnetic energy in space associated with a moving spherical charge is $\mu e^2 v^2/3a$ where μ is the magnetic permeability (equal to unity if the charge moves in a vacuum), e the charge, v the velocity, and a the radius. This then is equal to $Mv^2/2$ where M is the effective electromagnetic mass, which, it will be seen. equals $2e^2/3a$ (μ taken to be 1). This mass, being inversely proportional to the radius, becomes important when the latter is small.

However, this calculation of the electromagnetic mass is only valid providing the velocity of the charged particle is small compared with the velocity of propagation of electromagnetic waves, that is, with the velocity of light. Heaviside pointed out that as the velocity becomes great the tubes of force, as it were. crowd together and tend to set themselves at right angles to the line of motion. Lorentz showed that, in effect, a contraction of length takes place in the direction of motion. As a result of this the mass m_0 at zero velocity is increased to $m_0/(1-v^2/c^2)^{\frac{1}{2}}$ where v is the velocity of the charged particle and c the velocity of light. The electromagnetic mass of a charged particle thus increases with the velocity, but the effect only becomes noticeable when v is very large and indeed of the same order as the velocity of light, otherwise the factor $(1-v^2/c^2)$ is indistinguishable from unity: If the mass of the particle is entirely electromagnetic in origin if will increase in accordance with the Lorentz formula. It will be seen from Fig. 15.2 that Kaufmann's results show that the mass of the β -particle increases with velocity, but the accuracy of the measurements was not great enough to prove whether the Lorentz formula was obeyed or not.

Increase of β-particle mass with velocity

In order to examine the variation of mass with velocity more precise measurements were made by Bucherer whose experimental method is shown in Fig. 15.3a. The source of β -particles, radium fluoride, was placed at A, the centre of a parallel plate condenser. The condenser plates P_1 , P_2 , 8 cms. in diameter and 0.25 mm. apart, were maintained at a high potential difference. This circular condenser was surrounded by a coaxial cylindrical photographic film and the whole placed in a uniform magnetic field arranged so that the lines of force were parallel to the condenser plates. A β -particle will only succeed in escaping from between the plates P_1 , P_2 when the magnetic and electric forces acting upon it are equal and opposite, since under any other field conditions it will be deflected to one or other of the plates. The β -particles being

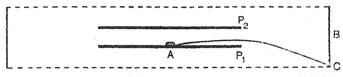


Fig. 15.3a

emitted along all the radii of the plates, can emerge at different angles to the magnetic lines of force, particles with different velocity succeeding in escaping in each direction. Consider a particle with velocity v moving along a radius at an angle θ to the direction of the magnetic lines of force. In order to escape the forces acting on it must be equal and opposite so that

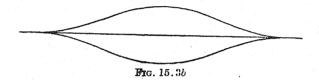
$$Xe = Hev \sin \theta$$
, i.e. $v = X/H \sin \theta$.

From this it follows that only particles with the velocity v will emerge along the direction θ . If any particle with velocity other than v starts out from A in the direction θ it will be deflected on to one of the plates and will not emerge. After leaving the plates each β -particle is now subjected to deflection by the magnetic field alone and traverses a circular path of radius $R = mv/He \sin \theta$. Thus from the deflection BC and the distance P_1B , R can be calculated for any given value of θ . Since the velocity is obtained from the value of θ , e/m can be

directly determined. By reversing X and H a symmetrical pattern is obtained upon the photographic film, which, when opened out has the appearance shown in Fig. 15.3b. Each pair of points gives a value for e/m for a given v.

Measurements by this method have been carried out using particles with velocities up to 0.8 that of the velocity of light. At this speed the increase in mass to be expected from the Lorentz formula is 66 per cent. that of the mass at slow speeds. The measurements have completely verified the law of increase predicted by theory.

It may be noticed that if it be assumed that the mass of an electron is *entirely* electromagnetic, which the mass variation at first appears to suggest, then the formula $m=2e^2/3a$ gives an estimate of the radius of the electron when it is at rest, or



moving slowly. The value of a can be obtained by simple calculation from the e/m measured for slow-moving electrons, and the value of e. Since

$$e = 4.80 \times 10^{-10}$$
 e.s.u. $= 1.60 \times 10^{-20}$ e.m.u.

and $e/m = 1.76 \times 10^7$ e.m.u. then $m = 9.1 \times 10^{-28}$ gms.

Substituting these values gives

$$\begin{array}{c} 9 \cdot 1 \times 10^{-28} = 2(1 \cdot 6 \times 10^{-20})^2 / 3a \\ \text{giving} & a = 1 \cdot 87 \times 10^{-13} \text{ cms.} \end{array}$$

This value is only an estimate but is proved by other work to be of the correct order.

In deducing this value for the electronic radius it has been assumed that the mass is *entirely* electromagnetic. The fact that the mass variation obeys Lorentz's formula was at first taken to confirm this. However, it can be shown from Einstein's Special Theory of Relativity that all moving mass, whatever its nature, obeys the law $m=m_0/(1-v^2/c^2)$. A

moving charge is only a special case of this more generalised law which applies to all matter. This being so, there is no valid basis for assuming that all the electron mass is entirely electromagnetic in origin, because even a residual gravitational mass would increase at the same rate as the electrical mass. It follows from this that the above deduction of the electronic radius is not necessarily rigid, and, in fact, the concept of the radius of the electron in modern theory is by no means a simple one.

The β-ray spectra

It has already been pointed out that the β -rays are inhomogeneous, consisting of particles with different velocities. This

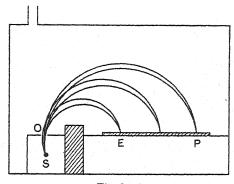


Fig. 15.4

lack of homogeneity is best demonstrated by the focusing method of Rutherford and Robinson which is illustrated in Fig. 15.4. The radioactive material which is the source of the β -radiation is placed upon a fine wire at S. The β -rays pass through O a fairly wide slit which is in the same plane as a photographic plate EP and vertically above S. The apparatus is evacuated and placed in a magnetic field with the lines of force arranged perpendicular to the plane of the paper. Those β -particles which have the same velocity (strictly speaking those which have the same momentum) are deflected through circular paths with the same radius and thus focus to an approximate point. The result is the production of a line

with a very sharp edge at the side nearer to P and this can be measured with great accuracy. A thick block of lead acts as a shield for the photographic plate, cutting off any direct radiations from S.

By measuring OS and the distances of the focal points from O, the various radii of curvature can be calculated. Since for a magnetic deflection mv = HeR, the momentum for each particle can be found. As the variation of mass with velocity obeys a known law, the energy of each particle can be calculated from the value of the momentum. It is usual to describe this energy in electron volt (e-volts) and not in ergs.

The origin of the β -ray spectra

Owing to the high velocities encountered it was initially considered that the particles constituting the β -ray spectra were electron groups sent out from the nucleus during a disintegration. This was, however, disproved by Rutherford, Robinson, and Rawlinson. It was known at the time that y-rays can eject very high speed secondary electrons from atoms when they are absorbed. It was suspected that this might be a possible cause of the observed β -ray spectra. The apparatus used to test this point was the same as that shown in Fig. 15.4, but the fine wire source S was replaced by a radon tube around which could be wrapped various thin metal foils. The γ-rays emitted by the source ejected secondary β-rays in their passage through the foils. The foil thicknesses were sufficient to stop most of the primary electrons and to straggle out the remainder. No definite lines due to the primary rays could appear on the plate. In spite of this a definite line spectrum was obtained. This could only have arisen from secondary electrons ejected from the foil by the γ -rays, suggesting that the β -ray line spectra themselves have a secondary origin.

Intensity measurements of the natural β -ray spectra were thereupon undertaken by Chadwick and it became apparent that the line spectrum only constituted a small fraction of the total β -ray emission. The main emission was in fact not in the form of lines, but was a continuum which could be identified as being the true nuclear disintegration radiation by counting the

number of particles given off from a known number of disintegrating atoms. Rutherford suggested that when a radioactive nucleus emits y-rays part of this radiation is absorbed by the outer electrons of the same atom, and as a result of this the secondary β -ray electrons are ejected. This process of internal conversion was subjected to direct test and proved to be correct. Ra B emits a γ-ray spectrum which is marked by three very strong lines, amongst others. When these are allowed to excite secondary β -particles from metals such as lead or platinum, etc., a secondary β-ray spectrum is formed which is similar to the natural β -ray spectrum of Ra B, only somewhat diffuse. Since the most intense of the lines in the excited spectra must be due to electrons ejected from the K shells of the atoms of lead, platinum, etc., it is possible to determine the absorption energies for the corresponding lines of the natural spectrum by comparing the β -particle energies. These absorption energies proved to be the K absorption energy of the radioactive atom, so it was proved that y-rays from a radioactive nucleus can be absorbed in the K shell of the same atom and thus liberate a secondary electron. Further, weaker lines due to quantum conversion in the L. M. N. . . . shells must also occur. This therefore accounts for the complexity of the observed β -ray spectra.

When a γ -ray is absorbed an amount of energy equal to the characteristic absorption of the state in question (K, L, M, . . .) is abstracted in order to eject the β -particle and the remaining surplus energy reappears as the energy of the particle. Thus the energy of the β -particle plus the known absorption energy for the characteristic state involved give the energy of the γ -ray. It has been possible to identify the shells from which the electrons are ejected in a very large number of the natural β -ray spectra which occur.

One point, however, has been left out in the above discussion. It is as follows. The K absorption energy effective in the internal conversion of any γ -ray is not that for the element from which the disintegration particles are emitted. The γ -ray is emitted after the disintegration particle and the nuclear charge, and hence the K absorption energy, has changed correspondingly (see p. 269).

The disintegration electrons

Although individual sharp lines appeared at first to be the most obvious characteristic of the β -ray magnetic spectrum, it was shown by Chadwick from an examination of the intensity distribution in the spectrum, that the whole region is covered by a diffuse continuous background, the integrated intensity of which is far in excess of that of the sum total of the individual lines. The method employed was an adaptation of the focus method already described, only instead of a photographic plate a Geiger counter or an ionisation chamber was used in the manner shown in Fig. 15.5. The radon source, Q, was in an evacuated box and the β -rays were deviated on to the

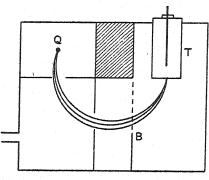


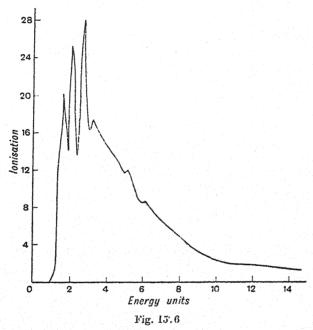
Fig. 15.5

counter T by a uniform magnetic field. The intensity of the field could be varied so that successive portions of the spectrum could be brought on to the counter. The movable screen B enabled the spectrum to be cut off so that the stray radiation effects could be separately determined and corrected for.

The result obtained in this experiment is shown in Fig. 15.6, from which it will be seen that the β -ray emission constitutes an extensive continuum upon which a few sharp lines are superposed. The sum total intensity of these lines is but a few per cent. of the total emission. All β -ray continuous emission spectra are found to be similar in their general shape and are characterised by a maximum and a definite sharp upper limit, the positions of which vary with different atoms.

In order to decide as to the origin of the respective line and

continuous spectra it is necessary to measure the total number of β -particles emitted. From the changes produced in atomic number by radioactive transformations it is known that in a β -ray disintegration each active atom gives off only one β -particle. Thus from a known quantity of active material, for which the rate of transformation has been determined, the number of β -particles given off per second can be calculated. If the β -radiation contains more particles than this, it can be



safely assumed that the excess is due to secondary β -particles excited by any of the radiations given off.

A reliable method, free from errors due to secondary electrons has been devised by Gurney for carrying out this measurement. The experimental arrangement is similar to that in Fig. 15.5, only the counter is replaced by a Faraday cylinder. The field is altered and the whole spectrum explored. Gurney found that each disintegrating Ra B atom emits approximately 1.25 electrons and each Ra C atom approximately 1.05 electrons. Theoretically one expects an emission of one electron per atom so that the excess must be attributed to the secondary electrons liberated by the γ -rays. In order

to test this point Emeleus made a count for the emission from Ra E which has no γ -radiation. Both the α - and β -rays emitted by a source of Ra(E+F) were allowed to fall upon a sensitive counter and the number of particles was determined. The β -rays were then deflected by a magnetic field and thus the contribution made by the α -particles was found separately. The numbers of α - and β -emitting atoms present at any time were known in terms of the disintegration constants of the two bodies. The result found was $1 \cdot 1 \pm 0 \cdot 1$ β -particles emitted per atom of Ra E.

The β-ray continuum and the neutrino

The evidence obtained from the study of distinct α -particle groups and that to be given later obtained from γ -ray studies, shows that in the nucleus excited quantum levels exist. It has been found difficult to reconcile this with the fact that the disintegration electrons form a continuum. A difficulty arising in connection with β -ray emission concerns the question of the conservation of angular momentum. Every atomic nucleus has an angular momentum associated with it (the nuclear spin) which is a half integral value of $h/2\pi$ when the atomic weight is odd and a whole integral value (or zero) when the atomic weight is even. The β -particle, like all electrons, has a spin of $(1/2).h/2\pi$ but when it leaves, say, an odd atomic weight nucleus, the remaining nucleus retains a half integral spin since its atomic weight remains odd. Obviously this destroys the conservation of angular momentum since a half cannot be taken from a half integral value and leave a half integral remainder. The same is true if the parent atom has an integral spin.

It can be shown that the very existence of the continuum, without a further hypothesis, implies that the conservation of energy is not maintained. Both these difficulties, that of conservation of energy and of angular momentum, have been accounted for by Pauli's postulation of a new particle, the neutrino. The conception of the neutrino has been developed in the following manner by Fermi. This particle is assumed to have zero mass, to be uncharged, and to have a spin of $(1/2) \cdot h/2\pi$. The emission of a β -particle is considered to be a double act, accompanied by a simultaneous emission of a

neutrino. The sum of the energies of the two particles is quantised and should indeed exhibit a fine structure similar to that of the α -particles. Since the theory postulates that the neutrino has no charge and zero mass (or extremely small mass compared with an electron) we have no means at present for detecting it. When the neutrino is emitted with zero energy the β -particle energy has a maximum value corresponding with the sharp end of the continuum. Theory shows that it is possible to deduce an estimate of the mass of the neutrino from the position of the maximum in the β -ray continuum and also from the shape of the end of the curve near the sharp limit. The application of this theory to the observed curves strongly suggest a zero rest mass for the neutrino.

The postulate that the neutrino and the β -particle both have a spin of $(1/2).h/2\pi$ leads to the retention of conservation of the angular momentum. It is well to realise that the existence of the particle is at present only hypothetical and that it has in fact been introduced to maintain the conservation laws. Chadwick and Lee, in an attempt to detect the neutrino, showed that it certainly forms less than one pair of ions in 150 kilometres of air at N.T.P., a fact which indicates the difficulty encountered in attempting to detect the particle.

The scattering of β-rays

Owing to the inhomogeneity of the β -radiations, the interpretation of scatter experiments, so fruitful in the case of α -particles, has proved to be difficult. In order to apply theoretical considerations it is necessary first to obtain a "monochromatic" beam of the rays by means of a preliminary sorting out with a magnet. This process at the same time deviates the rays away from undesirable α - and γ -radiations. As the angle through which a particle is scattered is inversely proportional to the square of its energy, β -rays, being so much less energetic, are scattered much more than α -ray particles. This means that the effective atomic scattering cross-section is relatively great, so that any observed scattering may be due to more than one atomic encounter. Furthermore, as the particles change their mass with change in velocity, complications arise in calculating effects to be expected.

The first reliable scatter experiments were carried out by Crowther. The β -rays from the source were first separated out by a magnetic field and a nearly homogeneous beam passed through a hole on to the absorbing screen which produced a scattered cone from the initial parallel beam. The particles were allowed to pass through a very thin screen and then into an isolated ionisation chamber where the current they produced was measured. Various circular stops were introduced in the beam, enabling the number of particles scattered between known angles to be determined. In order to avoid secondary electron effects, the whole apparatus was enclosed in an evacuated chamber. Crowther's results can only be interpreted if multiple scattering be assumed to take place, a fact that can be easily verified by the cloud chamber.

For single scatter to hold with a foil of given thickness, theory shows that a certain angle of scattering must be exceeded. This condition has been achieved by Chadwick and Mercier using a modification of Chadwick's method of studying α -particle scattering. The results found for the heavy elements agree with those predicted by theory. In the light elements the outer orbital electrons as well as the nucleus contribute to β -particle scattering. Abrupt changes in direction in the observed cloud tracks of β -particles, sometimes through more than a right angle, are direct cases of single scattering.

The absorption of β -ray particles by matter

Owing to the fact that the β -particles are scattered with ease, both the theoretical and experimental treatment of absorption are difficult. The ranges of the particles are large (about 100 times those of α -particles) and the straggle effects are great. The scattering effect is so large that with gold foils, for instance, as much as 50 per cent. of a β -ray beam is diffusely "reflected" by multiple scattering. Experiments on absorption have been carried out with an apparatus similar to that in Fig. 15.4, the source being wrapped in different thicknesses of the absorbing material. The curves found when the rays from a given source are absorbed respectively by paper, aluminium, tin, and platinum, etc., are similar in shape and

differences can be attributed to the different scatter effects which increase with the atomic weight. The end point in each curve is a measure of the β -particle range. This range is found to increase rapidly with the speed of the particles.

When the equivalent ranges of β -particles with different velocities (expressed in terms of mass per square centimetre) are determined in different absorbing media, they are found to fall upon a single smooth curve. From this it can be concluded that, to a first approximation, the particle range is independent of the nature of the absorber. This is only partially true however. The absorption of the slower moving β -particles can be studied from the lengths of the tracks in a cloud chamber.

Reduction of β-particle velocities by matter

The reduction in β -ray velocity produced by the passage of the particles through matter can be measured by the magnetic focus method. A single spectrum line can be isolated and the velocities of the particles constituting this line determined before and after absorption by measuring the radius of curvature in the magnetic field used to isolate the line. It is found that a fairly "monochromatic" line not only displaces (reduction in velocity) when the particles pass through an absorbing foil, but also broadens. For the purpose of measurement the photographic spectrum method of Fig. 15.4 is employed, the source being an activated wire bare for half its length and wrapped with foil on the other half. Two spectra are obtained, a displaced spectrum and a comparison spectrum, enabling the reduction in velocity for all the lines to be determined with one photograph. If the observed decreases are plotted against the weight per square centimetre of the absorbing material, smooth curves which approximate to straight lines are obtained.

The ionising power of the β-particles

The ionisation per particle can be directly measured by estimating the number of individual droplets (each formed round a single ion) in the cloud track of a β -particle. Such measurements show that in different gases the specific ionisation is approximately proportional to the number of electrons in the gas molecule, being, for instance, eight times as great in oxygen as in hydrogen. Below a critical velocity, as may be

expected, the β -particles fail to ionise altogether, but above this the ionisation rapidly increases to a maximum value. After reaching the maximum, the degree of ionisation falls with increasing speed to a constant value. The maximum ionisation occurs at an energy of about 1,000 e-volts, more than 1,000 ions being formed per centimetre in air at N.T.P. Very high speed β -particles produce only about fifty ions per centimetre under the same conditions.

REFERENCE

"Radiations from Radioactive Substances." E. Rutherford, J. Chadwick and C. D. Ellis. (1930.)

CHAPTER 16

THE Y-RAYS

The γ -rays

The y-radiations were discovered some time after the existence of the α - and β -rays had been established. The striking property they exhibit is their great power of penetrating even dense matter. The corresponding ionisation effects are relatively small, although the rays were first discovered by their ionisation properties. It was noticed that even a 30-cm. shield of iron failed to screen away some of the radiations from radioactive bodies. This penetrating radiation could not be deflected by the strongest available magnetic field, suggesting that it did not consist of charged particles. The radiation behaved like X-rays of extremely short wavelength. The high degree of penetration indicated that the wavelength was much smaller than that of even the hardest known X-rays. These y-rays were found to blacken a photographic plate, and if sufficiently intense to induce phosphorescence on a fluorescent screen.

Wavelengths of the \u03c4-rays

If the γ -rays are very short electromagnetic waves, it should be possible to measure their wavelengths by the use of a crystal as a diffraction grating precisely as has been done with X-rays. Rutherford and Andrade were the first to apply this method, using the simple and effective arrangement shown in Fig. 16.1. In these experiments a difficulty arises owing to the glancing angle being only of the order of a few degrees because of the extremely short wavelength. In Fig. 16.1, A is the source of γ -radiation and BC a large crystal of rock salt. At DE a photographic plate is set up. The rays from A strike BC at all angles of incidence, but for certain

angles reinforcement will take place, as in the case of the X-ray crystal method. These regions of reinforcement, which effectively constitute spectrum lines, will appear as blackenings upon the photographic plate. From the positions of the lines and the known spacings of the atomic layers in the rock-salt crystal, the wavelengths of the γ -rays can be calculated. If the radiations are not homogeneous each wavelength automatically selects out the correct angle for reflection.

Later workers have adapted the standard X-ray oscillating crystal method to γ -ray measurements. The radiations from a strong source are passed between a pair of plates about 50 cms. long. The plates are maintained at a potential difference of several hundred volts so that any β -rays given off by

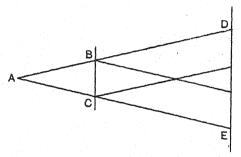


Fig. 16.1

the source will be deviated away. When a strong γ -ray source is employed the distance between it and the photographic plate can be increased to about 150 cms., enabling a higher dispersion to be attained. A large rock-salt crystal is slowly oscillated (2° in 24 hours). Some sources have more than 40 lines in the emitted γ -radiation. The observed wavelengths vary widely, those from Ra B and Ra C extending from 1,323 to 16 X.U. (The X.U. or X unit, is one-thousandth part of an angström, that is, 1.10^{-11} cms.)

There are four distinct methods by which the γ -radiations can be studied in general, namely, crystal reflection, absorption, excited secondary electron spectra and natural β -ray, spectra. The crystal reflection method has only a very limited application since it requires very strong sources. The absorption method gives very little information about wavelengths.

In the excited electron method, the γ -rays are allowed to fall upon a sheet of metal and the velocities of the ejected electrons measured, from which the wavelengths of the exciting γ -radiations can be calculated. This method also requires strong sources. As the natural β -ray spectra of radioactive bodies arise from the conversion of γ -radiation, an analysis of the β -ray spectrum enables the γ -ray wavelengths to be deduced. This method has a wide general application. The energy of a β -ray line is first determined, and then the line is identified with a given electron level of the atom. The known absorption energy for this level is then added to the observed energy of the β -particle coming from the level, and this gives the energy of the exciting γ -radiation $(\hbar \nu)$.

The absorption and scattering of y-radiations

The reduction in intensity which is produced when a beam of v-rays passes through matter is partly caused by scatter and partly by absorption. The true absorption will obey a logarithmic law if the radiation is monochromatic. Then the intensity, after passing through a thickness of material x, will become $E=E_0exp(-\mu x)$, μ being defined as the absorption coefficient. Most y-radiations are heterogeneous, but if increasing thicknesses of absorbing material are placed before a source, a stage is reached at which the softer rays are all cut out, only the hardest ray remaining. This is monochromatic, experiments showing that it exhibits true exponential absorp-This can be extrapolated back to zero and by subtraction the exponential of the next hardest component can be obtained. This process can be repeated in a manner similar to that used by Bragg in separating the effects due to α-particle ionisation in a complex case. By this means a series of absorption coefficients for the individual components of a complex yradiation can be calculated, but the method is clearly not very reliable, the effective resolving power being low. This process has, however, revealed the fact that the total y-radiation consists of hard, medium, and soft γ -rays given off by the nucleus, together with some characteristic X-rays of the outer electronic system of the radioactive atom. These characteristic X-rays arise from the interaction of the nuclear y-rays and the

electronic system of the atom through which they pass after leaving the nucleus.

When γ -rays are scattered by electrons, part of the scattered radiation suffers a wavelength change in accordance with the Compton effect. This complicates the investigations. Many theoretical attempts have been made to deduce a law of scattering, that of Klein and Nishina being most generally applicable. All the proposed theories assume that each electron in an atom has the same scattering effect, so that a scatter coefficient per electron is calculated. The scattering for the whole atom is taken to be Z times this, Z being the atomic number.

The determination of absolute values of absorption coefficients is difficult, due in part to the non-monochromatic nature of the radiations and in part to the fact that the actual ionisation depends not only upon the intensity but upon the wavelength too. Observation proves that the reduction produced in intensity by the passage of the rays through the light elements is almost entirely due to scatter, true photoelectric absorption of a γ -ray quantum being very infrequent. For the heavy atoms the photoelectric absorption per electron can be measured by subtracting the scattering for each electron (as determined with the light elements) from the total observed reduction in intensity.

Experiments on the direction of the scattered radiations show that the scattering predominates in the forward direction. Further, in accordance with the Compton effect laws, the greater the scatter angle the softer is the scattered radiation.

The γ -rays are not only spontaneously produced by the radioactive bodies but can also be excited by impact. When β -rays strike atoms, a continuous γ -radiation is emitted, the amount increasing with the atomic weight of the target. α -particles are also able to excite γ -rays by nuclear collision.

The energy given out in the form of γ -rays in the radioactive transformations of radon and its short-lived products is only about 7 per cent. of the total energy involved. This was proved by Rutherford and Robinson who measured the heat produced by the absorption of the α - and β -rays, by means of a *thin*-walled calorimeter which, being practically transparent

to the γ -radiations, allowed these to pass out. The heat measurements were then repeated with a *thick*-walled calorimeter which absorbed a known fraction of the γ -rays.

The act of emission of γ -rays

The y-rays are usually emitted by those atoms which also exhibit β -ray disintegration and it is important to determine whether the act of emission of the y-radiation occurs before or after the disintegration electron has left the nucleus. One method of deciding this is as follows. In a β -ray spectrum two groups of lines are identified as arising, say, from the K and the L levels of the outer electrons. The difference of the energies of these lines is the difference between the absorption energies of the K and L levels. However, this difference between K and L absorption energies can be determined for any atom from the X-ray spectrum, which shows that the K-L difference is not the same for the parent atom and the atom produced by disintegration. Measurements on the β -ray spectrum show that the γ -ray (which produces the β -ray spectrum by conversion) is only emitted after the parent nucleus has changed. i.e. after the disintegration electron has been sent out.

A different method of investigation which leads to the same result is afforded by the study of the secondary X-rays which accompany a disintegration according to the following mechanism. The β -ray line spectrum is produced by the ejection of outer electrons because of the internal conversion of the y-rays in the K, L, M, . . . states. It follows that an electron can fall back into one of these empty states after the β-particle has left, and in doing so a line of the characteristic X-ray spectrum is emitted. If the wavelength of this line is measured, the atomic number of the emitting atom can be determined. Rutherford and Wooster applied this method to the L lines of the X-ray spectrum of Ra B, using the crystal method for the wavelength determination, and found that the atomic number to which the lines correspond was 83. Since Ra B disintegrates from an atom of atomic number 82 to one of atomic number 83 it was proved that the y-ray was emitted after the disintegration.

The fact that the γ -rays are given off after the disintegration electron has been ejected implies that the loss of the electron has left the remaining nucleus in an excited state with a surplus of energy above that required for equilibrium. After a very short, but finite, interval, the nucleus returns to a more stable condition and in doing so emits its surplus of energy in the form of a quantum of γ -radiation. Ellis first suggested that the different frequencies in the γ -radiation given off by a radioactive material can be accounted for on the supposition of a number of excited energy levels in the nucleus.

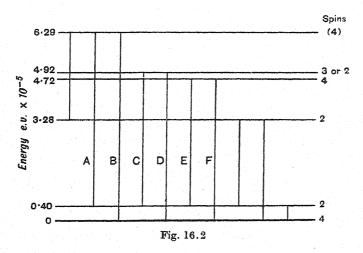
Nuclear energy levels

The first key to the study of nuclear excitation levels lay in the simple additive relationships found in the measured frequencies of the γ -rays emitted by some of the radioactive bodies. These relationships imply the existence of transitions between excited levels within the nucleus, a fact which can also be deduced from the fine structures which have been observed in the α -particles. Th C" affords an excellent example of the types of excitation levels met with. The β disintegration of Th B results in Th C which then emits an α -particle leaving behind Th C", which has a γ activity. The α -particles given off by Th C exhibit a fine structure, and on the basis of this, together with the additive relationships existing in the Th C" γ -ray frequencies, a group of nuclear excitation levels can be built up.

The excitation levels deduced for Th C" are shown in Fig. 16.2. The levels, and their energy values, are shown by the horizontal lines, the vertical lines representing the γ-ray transitions between them. A transition takes place when the nucleus readjusts itself from one excited state to another (the normal state being included amongst these). It is evident that a number of numerical relationships will occur, such as, for instance, the energy difference between the lines A—B equals that between C—D and between E—F, whilst that between C—E equals that between D—F, etc. This arrangement of levels is analogous to the term schemes found in optical and X-ray spectra. It has been found possible to link up the observations with theoretical deductions concerning the spins

of the nuclei, and it can be shown that the spins associated with the respective excitation levels introduce a selection rule. As a result of this only certain transitions are allowed, as will be gathered from the diagram of the Th C" levels.

It will be seen that the ten observed γ -ray energies require only six excitation levels in order to be fitted into a term scheme. These same levels successfully explain the fine structure of the associated α -particles also. Ellis has shown that the fifty-eight observed γ -rays of Ra C' can be accounted for as transitions between twenty-four excitation levels. In all



cases the excitation levels within the nucleus are quantised states, and the reversion of the nucleus back to the normal state with the accompanying radiation of a γ -ray quantum, is analogous to the emission of an optical or X-ray spectrum line by an excited atom. The existence of these nuclear levels has an obvious important bearing upon theories of nuclear structure.

General effects of a B-particle disintegration

The secondary effects which occur when a β -particle is emitted from a nucleus and this is followed by γ -radiation can now be summarised. In general the nuclear β -particles leave in the form of a *continuum*, that is, they have varying velocities. This fact has been accounted for by assuming the existence of

the neutrino. The nucleus when it is left in an unstable condition after the departure of the β -particle, reverts back to a more stable state, the surplus energy being radiated as α γ -ray. This it may do in a number of ways and it follows that the total γ -ray emission from a large group of atoms is not monochromatic, but consists of a number of distinct wavelengths related to the excitation energies.

Some of the y-rays, in their passage through the outer electron system of the atom which has radiated them, eject high-speed electrons from the K, L, M, etc., shells and these constitute the electrons giving rise to the β -ray line spectrum. As the y-rays have each a definite energy, and as a fixed amount of energy (the absorption energy) is required to liberate each K, L, M, etc., electron, the energy of each secondary electron is definite, being the difference between that of the incident y-ray and that of absorption. It is this fact which results in a definite line spectrum and not a continuum. The atom having emitted an electron (or perhaps a number of electrons) from the K, L, M, etc., shells is now in a state enabling it to radiate a line of its characteristic X-ray spectrum. A further minor complication can arise from the possible internal conversion of this characteristic X-ray spectrum within the atom, the result being the production of a weak secondary β-ray line spectrum, the energies of which are considerably lower than those due to the primary γ -rays.

REFERENCE

[&]quot;Radiations from Radioactive Substances." E. Rutherford, J. Chadwick and C. D. Ellis. (1930.)

CHAPTER 17

COSMIC RADIATION: THE POSITRON AND THE MESON

Historical

. The study of cosmic radiation has proved to be of immense scientific interest and importance, having added considerably to our knowledge of atomic nuclear structure and being instrumental in the discovery of two new particles, fundamental in the same sense as electrons, neutrons, and protons. origin of the study of this form of radiation goes back to early observations made in 1900 by Elster and Geitel and by Wilson upon the residual leaks of electroscopes. These investigators noticed that electroscopes always exhibit a small residual leak, no matter how excellent the insulation might be. The amount of leak was considerably reduced when the electroscopes were shielded with metal plates, proving that the larger portion of the leak is due to some sort of external radiation. Experiments carried out over deep lakes showed that most of this external radiation came from radioactive impurities in the earth.

These early investigations remained undeveloped until 1911, when Hess pointed out that the leak-producing radiation should diminish considerably on taking an electroscope to a great height in a balloon if all the radiation came from the earth. To test this crucial point he took ionisation chambers to a height of 5,000 metres in balloons. As expected, the ionisation at first began to decrease, but at 2,000 metres it reached a minimum in intensity and then started to increase steadily with height. To account for this increase it is necessary to assume that some form of penetrating radiation comes to the earth from outside. It is partially absorbed by the atmosphere and is thus more intense at higher altitudes. Hess further noticed

that the intensity of the oncoming radiation was the same during the day and night, and also during a solar eclipse. This suggested that the sun was not the source, and as the radiation appeared to be coming uniformly from outer space it was called "cosmic radiation."

The height to which a manned balloon can ascend is limited to about 18 Kms., even when special precautions are taken, hence, to extend observations, Regener instigated and developed the sounding balloon method of investigation. Automatically recording instruments, specially designed to withstand shock and to be as light as possible, are sent up in small sounding balloons. Quite a considerable fraction of these are recovered and by this means investigations have been carried up to a height of 30 Kms. All the observations prove that a steady stream of ionising radiation is pouring on to the earth from outer space.

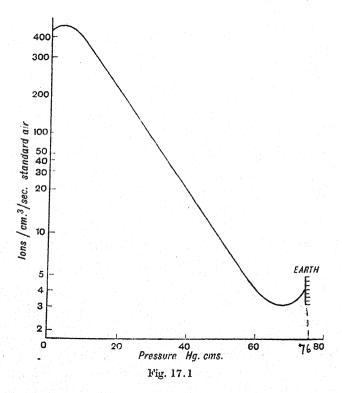
The absorption of cosmic radiation

Many investigators have contributed to the measurement of the absorption of cosmic rays by the atmosphere. Representative results are shown in Fig. 17.1 which illustrates the variation of ionisation with atmospheric pressure. As at a height of 30 Kms. the pressure is only about I per cent. that at sea level, in effect most of the atmosphere has been investigated. The ionisation reaches a maximum and then begins to fall.

Alongside with investigations of absorption carried out at great heights, experiments have also been made under relatively great depths of water. Regener measured the absorption of cosmic rays in water by lowering strongly built ionisation chambers in Lake Constance to 280 metres below the surface. At this depth the ionisation is so small that the chamber must be made as sensitive as possible. The bigger the volume of the chamber the greater is the chance of being struck by ionising radiations. Regener's ionisation chamber had a capacity of 39 litres and was filled with carbon dioxide at a pressure of 30 atmospheres, since the more atoms present per cubic centimetre in the chamber the greater is the ionisation produced by each cosmic ray.

The observed intensity at the bottom of Lake Constance is

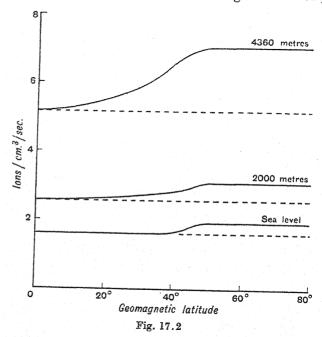
only 1 per cent. that at the surface, and about 10,000 times weaker than that at 16 Kms. above sea-level. More recently, Clay and co-workers have extended the water absorption measurements down to 440 metres and work in mines at depths equivalent to 1,000 metres of water, still show traces of radiation. The striking penetrating power of the radiation is evident when the degree of absorption is compared with that of very hard γ -rays. The Th C" γ -rays, with an energy



of 2.6×10^6 e.v. are reduced to 0.5 per cent. of their initial intensity after passing through 1.5 metres of water, yet cosmic radiation, after passing through 280 metres of water, has still an intensity 1 per cent. that at the surface. The penetrating power is so great that it is necessary to assume that some of the rays have energies exceeding 10^{11} e.v. Other evidence places the energy still higher.

The geographical distribution of cosmic rays

It was only after many expeditions had been made to various parts of the world that the existence of variation of cosmic ray intensity with geographical latitude was established. The investigations of Compton and his co-workers are typical, illustrating the amount of work that has gone into unravelling the effect. In 1933 Compton organised ten expeditions, each of which measured the radiation at a large number of places



between the geographical latitudes 46° S. and 78° N. At each observation station measurements were made at heights between sea-level and 6,000 metres. Some of the combined results from all the expeditions are shown in Fig. 17.2. Each of the three curves exhibits the variation of cosmic ray intensity with geomagnetic latitude (the latitude referred to the magnetic equator). The lower curve is for sea-level, the middle for measurements made at a height of 2,000 metres, and the upper for 4,360 metres. At all altitudes the ionisation remains constant from the pole down to 49°, at which point

there is a sudden drop in intensity, and the decrease continues to the equator. At sea-level the drop amounts to 12 per cent. but at higher altitudes it is more than 40 per cent.

In the expeditions just described free sounding balloons were employed. In addition to this, investigations in the stratosphere have been made by manned balloons up to a height of 18 Kms. In this type of investigation, first undertaken by Piccard, the observers must be encased in a sealed gondola. In one flight Cosyns drifted across Europe at a height of 12 Kms. and found the cosmic ray intensity to remain constant down to 49° N., after which it dropped suddenly. These measurements confirm those made by the sounding balloons.

Variation of cosmic rays with time

The absorption curve for cosmic radiation is not a simple exponential, implying therefore that the radiation is not of one type but mixed. It is possible, by the aid of screens of lead about 10 cms. thick, to cut out the softer incident rays leaving behind a component which is extremely penetrating indeed. By this means the two types of rays can be separately studied. Experiments extending over many years show that there is a small diurnal variation in the softer component, the intensity at midday being a few per thousand higher than during the rest of the day. On the other hand, the intensity of the penetrating component is practically constant. The constancy of these rays with time strongly suggests that the radiation does not come from the sun but is arriving uniformly on the earth from all directions.

The uniformity of incidence of the radiation introduces certain difficulties regarding its origin. At first sight it appears that the stars in the galactic system cannot be the source since the distribution of the latter is far from uniform. The possibility still remains that they are produced there and are smeared out into uniformity by stellar magnetic fields in space. The origin of the radiation remains at present unknown.

The nature of cosmic rays

As long as the ionisation chamber alone was used, little progress was made in discovering the nature of cosmic radiation. Two new powerful methods were developed and with

these the advance in recent years has been remarkable. The first of these new methods utilises Geiger-Müller counters. two or more counters are connected so that a record is only made when all the counters in the set are simultaneously discharged, this arrangement, called a coincidence counter set. becomes a very powerful weapon for the study of cosmic radiation. Thus when two counters are set vertically above each other, some distance apart, only the ionising particles coming down in a vertical direction can be recorded. In effect a cosmic ray telescope has been constructed, and as it can be pointed in any direction, the angular incidence of the rays can be investigated. Ionisation chambers on the other hand can only record the total radiation incident in all directions. By various groupings and arrangements of multiple counter sets any particular desired solid angle of incidence can be studied. A great deal of extremely valuable information has been accumulated by sending up counter sets in sounding halloons

The investigations made show that the rays travel in nearly straight lines. By interposing absorbing screens between the counters and applying the theory of absorption for charged particles it can be proved that the energies of the incident corpuscles are very high. This point will be discussed later.

An important result derived from counter coincidence set observations was the discovery of the azimuth effect. By pointing counters to the east and the west it is found, particularly at high altitudes and near to the equator, that the rays arriving from the west are more frequent than those coming in from the east. As will be shown later, this fact is an indication that a large number of the incident particles are positively charged.

Although both the ionisation chamber and the counter method have yielded valuable information, the most powerful weapon for the study of cosmic rays is the Wilson expansion chamber, particularly when used in the method developed by Blackett. The passage of a high-speed ionising particle through an expansion chamber creates a track of ions, and if the expansion takes place a very short interval after the passage of the particle a fine line track of droplets is produced and can be photographed. Normally cosmic rays are passing

through any chamber at fairly regular intervals, and if random expansions are made an occasional cosmic ray track will be observed. A re-examination of some of the very early photographs taken by Wilson show cosmic ray tracks, which had not then been recognised as such.

Clearly the study of cosmic ray tracks by means of random expansions is inefficient, slow, and laborious. This difficulty has been overcome in a very elegant manner by Blackett with the apparatus now to be described.

The Blackett counter-controlled expansion chamber

The details of the apparatus are shown in Fig. 17.3. If very sharp tracks are required the piston producing the expan-

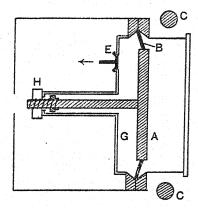


Fig. 17.3

sion must move very rapidly. It must therefore be very light and not subject to severe constraints. The piston A is made of aluminium alloy and together with the piston rod weighs only 280 gms. The observation chamber is made gas-tight by means of the flexible rubber diaphragm B. Both the chamber to the right of A and the space G to the left of A are filled with a gas at a pressure of 1.7 atmospheres, the valve E being closed. The amount of expansion is controlled by altering the position of the nut H which limits the stroke of the piston. The valve E is very light and moves back at high speed when released by a trigger which is pulled automatically in a manner to be described later. When the trigger is released the excess pressure of 0.7 atmospheres drives back the valve E and the

air under the piston rushes out. With this arrangement the piston takes only 0.005 seconds to complete its stroke.

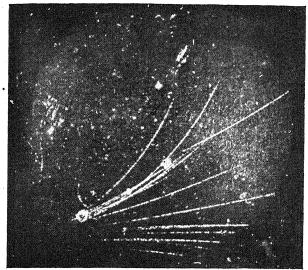
The details discussed above are concerned only with improving the definition, the main novel departure lies in the use of the counters. CC are Geiger counters placed directly above and below the expansion chamber. They are connected up so that a discharge only takes place when an ionising particle passes through both. Such a particle must also pass straight through the expansion chamber, being right in the field of view. By a system of amplifiers the current registered in the counters by the passage of a particle is made to operate a relay which pulls the trigger and sets the piston into motion. At the same time the chamber is illuminated and the resulting track photographed. Measurement shows that the expansion is completed within 0.01 second after the arrival of a particle. By this means 75 per cent. of the expansions which take place are actually due to cosmic ray particles. is clear that this beautiful method is rapid and economical.

In order to measure the energies of the particles passing through the chamber the latter is maintained in a very strong magnetic field. By means of a large electromagnet weighing 11 tons a uniform field of 14,000 gauss can be maintained over the whole area of the expansion chamber. Some of the incident particles are so energetic that even this high field fails to deflect them into an observable curvature. Special precautions require to be taken to avoid distortions in the photographic lenses or in the tracks (because of gas currents). The tracks reproduced in Plate VA are those observed with a field of 14,000 gauss imposed upon particles with energies 5×10^8 and 9×10^9 e.v. respectively. Special optical methods have been devised by Blackett for measuring the very slight curvatures encountered.

The effect of the magnetic field of the earth upon cosmic rays

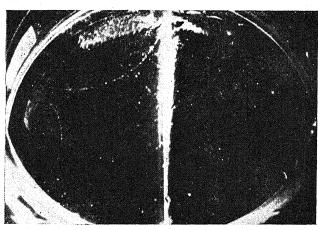
When high-speed electrically charged particles enter the magnetic field of the earth, they suffer deflection. The theory of this effect has long been studied with reference to the electrons given off by the sun and the manner in which the auroras are produced. The extreme penetrating power of cosmic ray particles is proof of the high energies involved. It

PLATE V

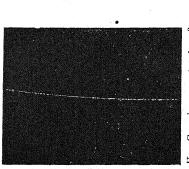


Vc. Cosmic ray shower.

(All after Blackett.)



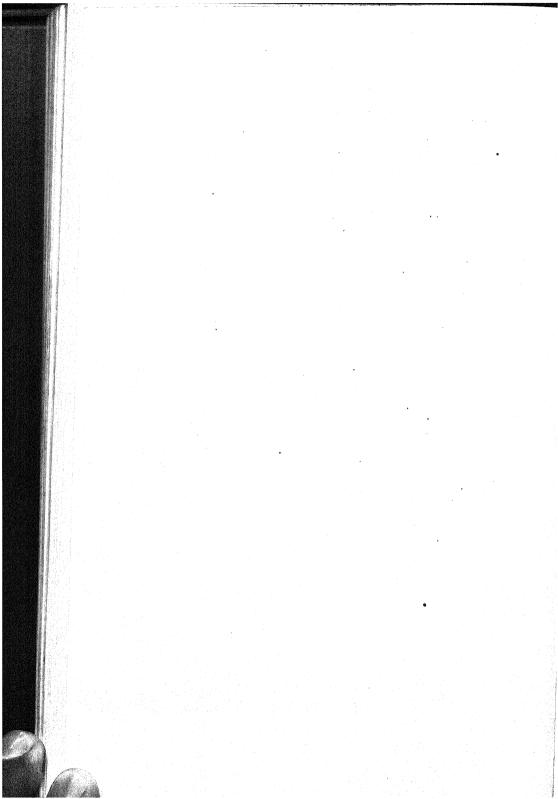
VB. Production of an electron-positron pair.



Va. Cosmic ray track of energy 5×10^8 volts. Magnetic field 14,000 gauss.



VA. Cosmic ray track of energy 9×10^9 volts.



COSMIC RADIATION: THE POSITRON AND THE MESON 28

is possible to calculate the orbits of high-speed particles in the magnetic field of the earth. The calculations made for the aurora by Störmer were first applied to cosmic ray phenomena by Rossi, and Störmer, then developed in detail by Lemaitre and Vallarta with the following results.

All particles with energies exceeding 6×10^{10} e.v. can effectively overcome the influence of the earth field and can reach all parts of the surface of the earth, no matter what be the direction of incidence. Particles with less energy than this are deviated so that there is a forbidden zone of latitudes, say between $\lambda_1 N$ and $\lambda_1 S$, which they cannot reach at all, and a second zone in each magnetic hemisphere, say between λ_1 and λ_2 , any point of which they can reach only in a given range of directions. For latitudes higher than λ_2 in each hemisphere the particles can reach the outer layers of the atmosphere in all directions. Clearly, as the energy of the particles decreases λ_1 and λ_2 increase.

It will be remembered that observations upon the distribution of the cosmic ray intensity over the surface of the earth show a very marked sudden decrease in intensity at latitude 49°. The theory shows that λ_2 for particles with energy 2.4×10^9 e.v. has this critical value. The reduction in intensity at this point implies that the incident radiation is mixed, consisting of particles with energy greater than 2.4 × 109 e.v., and the fact that some particles manage to arrive near the equator implies that energies exceeding 6×10^{10} e.v. also occur. This conclusion has been confirmed in a more precise manner by Blackett who has carried out extensive measurements upon the energies of cosmic ray particles by means of his counter controlled chamber in a magnetic field. It is now considered that owing to the sun's magnetic field, there are no incident primary particles with energy below 3×109 e.v. and the constancy of intensity at latitudes greater than 49°, at all altitudes, supports this general view.

The azimuth effect, already described, is produced also by the earth's magnetic field. If the incident particles consist of an equal mixture of positive and negative particles, there will be no azimuth effect, but if one type of charge predominates, the result of the direction of the earth's field is to produce an increase in intensity either from the east or the west according to whether there is an excess of negatively or positively charged primary particles respectively. The observed azimuth effect shows that the incident particles contain more positive than negative charges.

Although 2×10¹⁰ e.v. is the present limit for which energies can be measured in the expansion chamber, there is strong indirect evidence that some incident particles have energies a thousand times, and possibly a million times, greater than this. Such enormous energies are unknown in any other field of physics and are very much greater than those met with in radioactivity, which are of the order of 10⁶ e.v. For this reason the study of cosmic rays is of very great importance, since projectiles with energies of a very high order are now available for observational purposes.

The positron

A fundamental discovery arising from cosmic ray investigations was the detection in 1932, first by Anderson and later independently by Blackett of a new particle, the positron. It was noticed that among the cosmic ray tracks photographed with an expansion chamber in a magnetic field, were some identical with those produced by fast electrons, but deflected by the field in a direction opposite to that expected for a negatively charged particle. The tracks appeared to be produced by "positive electrons." These particles are now called positrons. Their existence is shown in a very striking manner by the interesting cosmic ray phenomenon known as cosmic ray showers, discovered by Blackett and Occhialini with the counter controlled chamber. A typical shower is shown in Plate Vc. It consists of a group of tracks originating from a point outside of the chamber, usually in the massive metal parts of the apparatus. Since the cloud chamber is in a magnetic field the tracks are curved. Each shower consists of an approximately equal number of positively and negatively charged particles, since the curvatures of about half are in an opposite sense to those of the rest. From the energies and amount of ionisation it can be concluded that the tracks are due to positive and negative electrons. It is clear that the shower must originate in a definite small region, a fact which decides the direction of motion of the shower particles. Hence the direction of

curvature unambiguously fixes the sign of the charge. Shower production will be discussed in more detail later.

After its discovery in showers it was found that there are other sources of positrons. When hard γ -rays are absorbed by matter the quantum of radiation is sometimes completely absorbed by an atom which simultaneously emits a positron and a negative electron. This phenomenon is called pair production, and if the absorption takes place in a cloud chamber maintained in a magnetic field a paired track is seen to come from the atom absorbing the quantum. Similar pairs are also sometimes seen in cosmic ray photographs. The Plate VB (reproduced by permission of Professor Blackett) is an example of pair production in a magnetic field.

The harder the γ -radiation the more effective is it in producing positrons. The positron yield also increases with the atomic number of the atoms which are struck. The positrons created by this process of conversion of γ -radiation have approximately the same energy, no matter what the absorber may be. This suggests that the pair production does not arise from a nuclear interaction but occurs in the Coulomb field, a fact in agreement with theoretical prediction.

Pair production by absorption of γ -radiation is a remarkable phenomenon, for here there is a complete transformation of a quantum of radiation into matter (electrons and positrons). The converse process, the annihilation of a positron and an electron which have come together has also been observed. In this case "annihilation radiation" is emitted and can be detected.

Many sources of positrons are available. When light elements are bombarded with α -rays, pair production sometimes results, due probably to the internal conversion of the γ -rays given off after the nuclear collision. The most copious supplies, however, can now be obtained from artificially produced radioactive substances, the properties of which will be discussed in a later chapter. In these only positrons, and not pairs, are ejected. These are emitted from the nuclei of the atoms concerned, not from their outer electron systems (inter-action with the Coulomb field) as is the case in pair production.

The mass of the positron

Strictly speaking the cloud tracks attributed to positrons only indicate the fact that the charge is positive and that the amount of ionisation is similar to that produced by an electron, being the same to within 20 per cent. The latter fact shows that the charges of the electron and positron are approximately equal. The mass is best estimated by observations made upon pairs produced by the absorption of γ -rays of known energy. Suppose that a y-ray of energy E creates two oppositely charged particles of masses m_{+} and m_{-} . At the instant of creation the particles are ejected with energies e+ and erespectively. It is known from the Relativity Theory that, if c is the velocity of light, an amount of energy equal to mc^2 is required to create a mass m. As the original γ -ray energy is used up in creating the particles, the surplus reappears as the kinetic energy. It follows then that $E = (m_+ + m_-)c^2 + e_+ + e_-$. Since the γ -ray energy E and the electron mass m_{\perp} are known, the positron mass m_+ can be calculated if the kinetic energies of the particles e_+ and e_- are measured. The particle energies can be derived from track curvatures in a cloud chamber maintained in a magnetic field, if values of charge and mass are assumed.

It is not always possible to measure both members of a pair with the same accuracy and in this case a slight modification of the method can be used. Suppose the positron and electron masses are equal, then by numerical substitution we get $(m_+ + m_-)c^2 = 1.01 \times 10^6$ e.v. If the 2.62×10^6 e.v. γ -radiation from Th C" is used to produce the positrons it follows simply that $e_+ + e_- = (2.62 \times 10^{6} - 1.01 \times 10^{6})$ e.v. =1.61×106 e.v. It is clear that the maximum energy a positron can acquire under these conditions occurs when the electron energy e_ is zero. Thus, if the masses of the electron and positron are identical, no positron will ever have an energy exceeding 1.61×106 e.v. Anderson first pointed out that these methods would give a value for the positron mass. Chadwick, Blackett, and Occhialini, who measured the energies of a large number of positrons, finally found that $m_{+}=(1.02\pm0.10)m_{-}$. Within experimental error it is therefore proved that the positron and electron are identical except as regards charge.

When a statistical distribution of the energies of the positrons

and electrons in pairs is carried out, it is found that for a number of cases the positron energy is greater than the electron energy. This is because the pair of particles is created in the neighbourhood of an atomic nucleus which is positively charged. If both the particles are generated with the same amount of kinetic energy, the positron, being repelled by the nuclear field, acquires an increased energy, whilst the electron, being attracted, has its kinetic energy reduced.

The nature of the positron

Some of the fundamental properties of the positron can be derived from the application of conservation principles. Since an electron and positron are created by the disappearance of a quantum which has no charge, it is clear from the fact that the charge is conserved that the electron and positron charges must be equal and opposite. The derivation of the positron mass from the principle of the conservation of energy has already been discussed. A further property can be derived from the conservation of angular momentum. It is known from spectroscopic and specific heat studies that the electron has an angular momentum equal to $(1/2) \cdot h/2\pi$. For angular momentum to be conserved the positron must have a spin numerically equal to this, for there is evidence that a radiation quantum cannot have a half integral spin associated with it. Hence the charge, mass, and spin of the positron are known.

The experimental detection of the positron has been a triumph for Dirac's theory of the electron, which had predicted the existence of the positron long before it was discovered. Dirac had applied the quantum mechanics to the derivation of the wave equations for an electron moving in an electric field. By taking into account the effects due to relativity a theoretical explanation for the electron spin was obtained. However, the complete solutions to Dirac's equations resulted in electron states with positive kinetic energy and negative kinetic energy. According to the theory, transitions between these states should take place, but since such transitions did not appear to be observed in nature, Dirac inferred that practically all of the predicted negative energy states are usually filled up.

He was able to show that if a "hole" occurred in one of these negative energy states, i.e. if one were unoccupied, this "hole" would correspond to an observable particle. Such a particle would have the same mass and spin as an electron but would have a positive charge. In fact, such a "hole" would be a

positron.

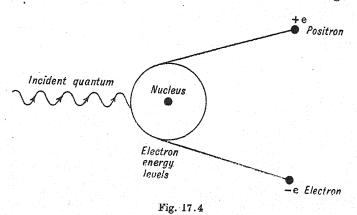
It is of interest to inquire why the positron remained undetected for so long. The answer was given by Dirac who showed that, owing to the fact that a positron tends to rush towards an electron with resulting annihilation of both, the mean life or existence time of a positron depends upon the number of electrons near to it. The concentration of the latter in matter is very high and as a result the mean life of positrons is very small. Calculation shows, for instance, that the mean life in water is only 10^{-10} second. It is the extremely short life of the positrons that explains why they can only be observed in a cloud chamber or with a Geiger counter immediately after being created. As soon as a positron approaches an electron the particles rush together and annihilate each other. They completely disappear and in doing so give out radiation with total energy $h\nu = \mathbf{E} = (m_+ + m_-)c^2 + e_+ + e_-$. This annihilation radiation has been detected when positrons are allowed to be absorbed by a sheet of metal.

The mechanism of pair production by the absorption of a quantum of radiation is as follows. Radiation can be transferred to electrons by one of three methods, photoelectric effect, Compton effect, and pair production. In the photoelectric effect the quantum is completely absorbed by the ordinary outer electronic system of an atom and its energy reappears as kinetic energy (minus work of emission). In the Compton effect the quantum transfers a certain proportion of its momentum to the electron, and passes on in the form of degraded radiation. In pair production, characteristic only of high energy quanta, an electron is lifted out of a negative energy state in a manner analogous to the ordinary photoelectric effect. This constitutes the negative member of the pair. However, a hole has been left behind in the negative energy states and, by Dirac's theory, this is in effect a positron left behind. Hence the ejection of an electron from a negative energy state automatically creates a positron at the same

COSMIC RADIATION: THE POSITRON AND THE MESON 287 time. This is the mechanism of pair production. It is

illustrated in Fig. 17.4.

It has been suggested that many of the incident cosmic ray particles are positrons. If this is true outer space must contain very many positrons. They can exist there since the negative



electron concentration in space is extremely low compared

The formation of cosmic ray showers

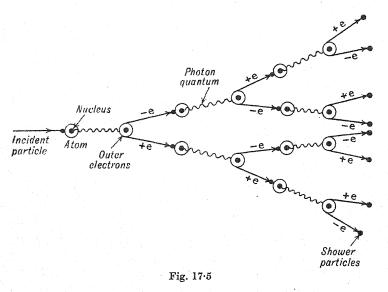
with that in matter.

Effects really due to cosmic ray showers had been noticed years before the showers were recognised, being shown by ionisation chambers at all altitudes. Occasionally an ionisation chamber would show a very sudden burst of ionisation due, as we now know, to a shower of cosmic ray particles. Showers can therefore be studied at sea-level with expansion chambers or at varying altitudes with ionisation chambers by recording the bursts.

It is now established that the majority of showers are produced by the cascade process illustrated in Fig. 17.5.

A very energetic cosmic ray particle strikes an atomic nucleus, losing a large amount of energy in the form of a photon. The principle of conservation of momentum requires that this photon shall be emitted in the forwards direction. After travelling a short distance, the photon collides with an atom and is absorbed by interaction with the Coulomb field near the nucleus leading to the process of pair production, which

is a photoelectric effect upon a virtual orbit of negative energy. The members of the pair are similar in nature to the incident particle (which was either an electron or a positron) but have roughly only half its energy. Each member of the pair is capable of liberating a quantum by collision with an atomic nucleus. Each photon so liberated can in turn produce a pair. Clearly after passing through sufficient matter the original incident particle is transformed into a group of



electrons and positrons approximately equal in number. (If every particle creates a photon the number will be exactly equal.)

The theory of this cascade process has been worked out in detail by Bhabha and Heitler, who have shown that a plate of lead 1 cm. thick suffices to produce a hundred particles in a shower formation. Such showers in lead plates can be studied in the laboratory by placing the plates either just outside of, or within, an expansion chamber. The low-energy particles observable at sea-level are largely secondary particles arising from showers created in the atmosphere. The effective thickness of the atmosphere is such that showers consisting of many thousands of particles, and extending over quite large areas, can be created, and have been observed.

Heisenberg has put forward the view that some showers are not cascades, but are created by an explosion process. According to his theory, an incident particle strikes a nucleus and thereby an energetic photon is liberated. This is absorbed by an atom, but instead of a single pair being formed, multiple pairs are created by a single operation. There is some evidence that this explosion mechanism possibly takes place.

By arranging counters in triangular patterns Rossi and other workers have considerably extended the study of cosmic ray showers.

The meson or heavy electron

The discovery of the heavy electron, now called the meson, has a striking similarity to the discovery of the positron, since both were predicted some time before being recognised and both were discovered during the process of cosmic ray investigations. The study of the penetrating component of cosmic radiation resulted in a number of difficulties. Calculation showed that the mean free path of a shower-producing electron is only 0.6 cm. in lead so that even very energetic electrons are quickly absorbed by the cascade process whilst traversing lead. However, at sea-level about 80 per cent. of the cosmic radiation consists of extremely penetrating particles of which 50 per cent. pass through 100 cms. of lead.

Clearly these penetrating particles are not electrons. Theory shows that the mean free path of penetration is proportional to the square of the mass of the penetrating particle (for low mass), from which it must be concluded that the penetrating component of cosmic radiation consists of particles of mass greater than that of the electron. To determine the mass of a cosmic ray particle it is necessary to measure both the curvature of the track formed in a magnetic field and also the number of ions produced per centimetre of path. It so happens that with very fast particles the amount of ionisation produced is almost independent of the mass, an electron and a proton being equally effective. Only when the particles have slowed down is there a measurable difference. Some twenty tracks of slowed down penetrating particles have been recorded and from these it has been proved that there exist particles,

either positively or negatively charged, with masses many times that of the electron. These particles have been named mesons (particles of intermediate mass). The first meson track was identified by Anderson in 1937.

Yukawa's theory of the origin of the meson

After the experimental discovery of the meson attention was drawn to the remarkable prediction of such a particle made by Yukawa in 1935. The history of the discovery of the meson is in fact almost identical with the history of the discovery of the positron for in both cases mathematical theory of an abstruse nature had predicted the particle, the theory in each case remaining more or less neglected, and after the experimental discovery of each of the particles, the earlier theory was recalled.

The Yukawa theory is a modification of the Fermi theory of B-decay for radioactive materials. An atomic nucleus contains in it neutrons and protons (perhaps bound together in a-particles). Neutrons are uncharged, hence any force operating between a neutron and neutron or a neutron and proton is not of the coulomb type. Experiments carried out upon the scattering of neutrons by protons show that forces operate between these particles within a range of 10-13 cms. Such a force had been postulated by Fermi to explain β -decay. Fermi supposed that the neutron and proton were different quantum states of the same fundamental particle and could transform one into the other by the quantum emission of a particle. Thus if a proton loses a positive electron it becomes a neutron, and if a neutron loses a negative electron it becomes a proton. A proton and neutron are bound together by what is called an "exchange force" arising because they are rapidly changing their identity by the exchange of the particles just postulated. The conception of exchange force is not new and has been used very successfully to explain chemical valency and various aspects of spectroscopy. The exchange can be represented as P←e→N where P and N are the proton and neutron and e is the exchanging particle. According to the Fermi theory, electrons or positrons jump to and fro and sometimes escape, and this constitutes the mechanism of β -decay, the decay period depending upon the probability of escaping.

It is possible by this theory to calculate decay periods and also the binding energy between the proton and the neutron. It is found that there is complete discrepancy between calculations based upon observed β -decay and calculations based upon binding energies derived from the mass defect of deuterium, etc. Yukawa thereupon attempted to solve the difficulty by modifying the theory in the following manner.

Yukawa proved mathematically that it was impossible to assume that the exchanging particles producing the exchange forces are electrons and positrons. Instead of this, theory demanded that the particle should have the charge of the electron, either positive or negative, and should have a mass approximately 100 times that of the electron. The exchange then is represented by $P^{\leftarrow\mu\to}N$ where μ is the particle with mass $\sim 137m$ (m is electron mass). This particle is probably to be identified with the meson observed in the expansion chamber.

The instability of the meson

The mass value assumed for the meson is adopted in order to give the correct binding-energy values. It remains to explain β-decay and to do this Yukawa makes a further, assumption, namely, he postulates that the meson can spontaneously disintegrate, being in this sense radioactive. If the meson is represented by μ the decay is shown by $\mu \rightarrow e$ +neutrino. If the original meson is positively charged (theory demands both types) the decay particle is a positron, if it is negatively charged, then the decay particle is an electron. The charge and spin are conserved, the spin of the meson being an integer. Since the electron mass is very much less than that of the meson the electron must be ejected with high energy equal to the loss in mass. Calculation shows that the electron energy will be about 4×10^7 volts. In order to explain β -decay the meson must have a natural decay period. When this is evaluated it is found to be about 10-6 second.

The particle postulated by Yukawa is more probably identified with the heavy electron found in cosmic rays, since the predicted instability has been confirmed experimentally. The penetrating component of cosmic rays consists largely of mesons, and if these decay in 10⁻⁶ second, they are clearly



produced in the atmosphere and do not come from outside. From certain anomalies in the absorption of this penetrating radiation a decay period, which is approximately 10⁻⁶ second, has been derived. Blackett has obtained a similar value from measurements made on the fluctuations of cosmic ray intensity produced by alteration in the atmospheric temperature. The values so calculated agree well with those predicted by Yukawa.

Properties of the meson

Measurements of the mass of the meson are relatively inaccurate, involving a high degree of uncertainty. Some twenty determinations have been made, the observed values varying from 39 to 570 times the electron mass. It may be that all mesons have not the same mass, but the evidence, taking into account the high errors of observation, is entirely inconclusive on this point. The detailed Yukawa theory requires the existence of positive, negative, and neutral mesons. The neutral particles have been called "neutrettos," having the same supposed mass as the other two but no charge. The symbols used for the particles are m^+ , m^- , m^0 , respectively.

The three particles are related to protons and neutrons according to the following reactions. A proton on emission of a positive meson becomes a neutron, a neutron on emission of a negative meson becomes a proton, and the emission of a neutretto is equivalent to the emission of a light quantum in the sense that the state, but not the nature, of the emitter changes, thus:

$$p \rightarrow n + m^+$$

$$n \rightarrow p + m^-$$

$$p - p + m^0$$

$$n \rightarrow n + m^0$$

With free protons or neutrons these processes cannot actually take place, no mesons being set free, as energy is needed for the creation of the meson mass and this is not available. However, in the immediate neighbourhood of the proton or neutron the meson has a virtual existence, being in effect continuously emitted and absorbed, but unable to escape. If the meson is not reabsorbed by the same heavier particle (proton)

the whole time, but is absorbed by a very nearby neutron, the latter will become a proton, and the original proton becomes a neutron. Effectively then the neutron and proton have interchanged places. Such conditions of closeness of particles occur in nuclei, and the resulting binding between the particles due to the change of the position of the meson is the exchange force. The neutretto is required by theory to account for the exchange forces needed to bind like particles such as protons and protons.

Since a free proton is for a part of its lifetime a neutron with a closely bound meson, the magnetic moment of a free proton is the resultant average of the true magnetic moment of the proton and the magnetic moment of the meson, taking account of the relative time spent in both states. The proton is only "dissociated" for a small fraction of its life, but as the meson is much lighter than the proton, the meson magnetic moment is considerably greater. Thus the net magnetic moment of a free proton should exceed that expected by simple theory. This is indeed observed. Further, in a strictly analogous manner a neutron is, for a fraction of its lifetime, dissociated into a proton and a negative meson. (This combination has a negative magnetic moment.) It follows that, although uncharged, a neutron will have a negative magnetic moment. This again is observed experimentally.

The penetrating portion of the cosmic radiation consists of free mesons, the disintegration of which has been observed in the cloud chamber by Williams. Their origin is a problem not yet explained. The theory of collision and absorption between atoms and mesons has been worked out on the basis of cosmic ray research results. When a meson collides with a neutron within the nucleus of an atom (clearly such events are frequent in the air) multiple processes can result. To take only a simple example, let a positive meson react with a neutron then either

$$m^{+}+n \rightarrow n+m^{+}+m^{-}+m^{+}$$

can represent what occurs, or possibly

$$m^+ + n \rightarrow n + m^+ + m^0$$

or, alternatively, we may get

$$m^{+} + n \rightarrow p + m^{-} + m^{+}$$

and so on, with more complex reactions involving more particles, all the result of one collision. Methods of calculation begin to break down and indeed Heisenberg is of the opinion that the quantum theory itself begins to break down and is no longer applicable. In effect the process under consideration, in cosmic ray phenomena, leads to an explosive shower as distinct from a cascade process shower.

REFERENCES

[&]quot;Cantor Lectures," The Royal Society of Arts. P. M. S. Blackett. (1937.)

"International Conference on Physics." P. M. S. Blackett. (1935.)

Nature. Vol. 142, p. 692 P. M. S. Blackett. (1938.)

CHAPTER 18

THE NUCLEUS OF THE ATOM

Introduction

The success attained in dealing theoretically with the electronic structure of the atom can be attributed to the fact that the constituent components (nucleus and electrons) are relatively so small compared with their distances apart that they can be treated as point charges. Exact calculations can therefore be made and theory can be checked. However, the problem of finding how the nuclei of atoms are constructed is more difficult, since it is known that nuclei are very small. The nuclear matter must be very compact and must also have some form of structural arrangement of packing of the constituent particles. The closeness of packing is the main difference between nuclear and extra nuclear structure. It is known that the forces acting between particles which are very close together differ from the force which acts when they are relatively far apart. These special forces give the nuclei their special properties.

Within recent years a great deal of progress has been made in developing a comprehensive theory of nuclear structure. Many independent lines of investigation have contributed to the problem. The experimental data have been derived from studies of mass spectra, the phenomena of radioactivity, the investigation of collision effects between atomic nuclei and high-speed radioactive or artificially accelerated particles, cosmic rays, the study of the spin properties by spectroscopic means, etc. Theoretical investigations have gone hand in hand with experiment, sometimes explaining observations sometimes predicting new properties, and as a result of the great deal of work that has been done and is still being carried out, a satisfactory theory of the nucleus is evolving.

Of primary importance are the whole number rule and the mass defects, already discussed. These indicate, firstly that the different nuclei are built up from integral numbers of simple units, secondly that the degree of binding of these units differs from atom to atom. The basic simple building stones of nuclear structure are the proton and the neutron. It is not yet known with certainty whether a-particles (helium nuclei) exist as real independent sub-units in nuclei, but there is a distinct possibility of this being the case. The a-particle is a stable arrangement of two protons and two neutrons. The fact that radioactive nuclei can eject a-particles strongly suggests that such particles exist as individuals in at least the heavy radioactive nuclei.

Gamow's theory of a-particle disintegration

The spontaneous emission of α -particles, characteristic of so many radioactive nuclei, has been explained by Gamow who has applied wave mechanics considerations to the problem of nuclear stability. This has led to an important advance in the conception of nuclear structure. Wave mechanics analysis, in general, shows that there is often a small but finite probability for the occurrence of events which are normally absolutely forbidden by classical mechanics. The event considered in this connection by Gamow is the spontaneous emission of an α -particle.

Gamow postulates the existence for a radioactive nucleus of a potential barrier which is a space in which the potential is so high, that, according to classical conceptions, an α -particle inside this barrier cannot escape. The distribution of potential within and around the nucleus is similar to that shown in Fig. 18.1. The centre of the nucleus is at O, the ordinate X representing the potential at any point distant r from the centre. If the nucleus is approached from some distance the inverse square law is obeyed, the potential curve CD being a rectangular hyperbola. As the centre of the nucleus is approached (r made small) the curve drops to form a potential well. If PM represents the kinetic energy of an α -particle (strictly, the kinetic energy per unit charge), then, according to the classical point of view, the particle can never escape

from HH. In the wave mechanics treatment the wave representing the a-particle is reflected to and fro within the barrier, behaving virtually as a stationary wave. The theory shows that part of this wave leaks out of the potential barrier. Physically this means that the equivalent particle has a small but finite probability of escape.

The probability of escape is a measure of the radioactive transformation constant, since the more probable the escape the shorter is the effective radioactive period. If the probability of escape is small the transformation constant is also

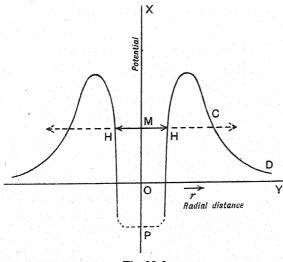


Fig. 18.1

small. The higher the position of M in the diagram, that is, the higher the kinetic energy of the a-particle, the greater is the probability of escaping, from which it follows that a high transformation constant will always be associated with long-range particles. The detailed application of this theory enables a relationship to be derived between range and transformation constant. This is found to be the same as the experimental Geiger-Nuttall law.

As the a-particle exists within the nucleus in the form of a stationary wave, there are different normal modes of vibration. That is to say, the equivalent particle can have different

distinct energy values. It is to be expected from this that nuclei can go into states with different degrees of excitation. As already discussed in connection with a-, β , and γ -ray emission, this is the case.

A particularly interesting application of the theory, and indirectly a confirmation of it, exists in connection with the experiments made by Rutherford and his co-workers on nuclear disintegration by means of accelerated particles. Although the study of natural radioactivity is of great value, more rapid progress in the understanding of nuclear processes has been made by the development of the technique of artificially disintegrating atomic nuclei. This will now be considered.

Artificial disintegration of nuclei by α -particles

It was long ago realised by Rutherford that the high-energy a-particles could be used as projectiles in an attempt to disintegrate atomic nuclei. When an a-particle is shot into a gas the chance of a collision between the particle and the nucleus of one of the gas atoms is very small, owing to the relatively small nuclear radius. However, in 1919 Rutherford succeeded in observing such collisions and thereby produced the first experimental transmutation of matter, the dream of the alchemist for centuries. Rutherford found that when a-particles are shot into nitrogen gas, high-speed protons are ejected from the nuclei of the nitrogen atoms. He followed up this investigation, examining twelve light elements and in each case a-particle impact caused the violent ejection of a proton. A special method was devised for the determination of E/M for the ejected particles and by this means it was proved conclusively that they were protons, and not nitrogen nuclei which had been "knocked on." The protons had surprisingly long ranges, those from aluminium going as far as 90 cms. in air.

The mechanism of the production of these long-range high-speed protons was demonstrated by Blackett in 1925 by means of expansion chamber observations. An example of this method of investigation is shown in Plate IVD. This shows the tracks of α -particles through nitrogen gas. The forked track, which is clearly visible, shows where an α -particle has struck a nitrogen nucleus. The long, thin track from the fork

is that of the ejected high-speed proton, the shorter thick branch being the track of the recoil atom. The a-particle enters the nitrogen nucleus and forms an unstable combination, which, after a short but finite interval, disintegrates, ejecting a proton. An isotope of oxygen, O^{17} , is left behind. This will be discussed in more detail later.

When calculations are made, taking into account the known energy of the a-particle and the probable height of the nuclear potential barrier, it is found that the energy of the particle in the case illustrated in Plate IVD is classically not sufficient to enable it to pass through the barrier. However, the Gamow wave mechanics theory indicates that there is always a finite probability of penetration because of the leakage of waves through the barrier. Nuclear disintegration can therefore be accomplished with the aid of energies considerably less than at first anticipated.

The neutron

The study of the disintegrations of light atomic nuclei by a-particles led to the discovery of the neutron, by Chadwick, in 1932. Already in 1920 Rutherford had predicted the existence of a nuclear particle which possessed mass without charge, but he had failed to detect it. In 1930 Bothe noticed that when beryllium was bombarded with a-particles no protons were ejected, as at first anticipated by analogy with nitrogen. Instead of this a penetrating radiation was emitted which he interpreted naturally as being a form of γ -radiation. This radiation was able to expel high-speed protons from paraffin wax. Chadwick showed that the number and range of these protons was such that the incident radiation could not be γ -rays, but consisted of uncharged particles with a mass approximately that of the proton. He called these uncharged particles neutrons.

Neutrons are now known to be constituents of every atomic nucleus (other than that of hydrogen which is a simple proton). The absence of charge enables the neutrons to penetrate matter very easily and also makes it *impossible to obtain tracks of neutrons in an expansion chamber*. Cases have, however, been observed where a neutron collides with an atom in an expansion chamber and at the point of collision tracks suddenly appear without any apparent origin.

Experiments made on the penetrating power of neutrons from beryllium show that 10 per cent. of the incident neutron radiation is able to pass through 30 cms. of lead. Certain light bodies like paraffin wax have, on the other hand, a high absorption coefficient to neutrons. When an element is excited into neutron emission by impact with a-particles, neutron groups with distinct energies are emitted. These different groups are related to different excitation levels within the nuclei.

It is convenient to express nuclear disintegrations in the form of a nuclear reaction equation, similar to the reaction equations used in chemistry. A typical equation, given below, is that for the emission of a neutron by the collision between a boron nucleus and an α -particle. This is represented by

$$^{11}_{5}B + ^{4}_{2}He = ^{14}_{7}N + ^{1}_{0}n.$$

The left-hand upper superscript gives the mass of each particle and the left-hand lower superscript the charge. A boron nucleus of mass 11 captures a helium nucleus of mass 4, the combination breaking up into ordinary nitrogen and a neutron. Both nuclear mass and nuclear charge are conserved in the reaction. Any mass apparently lost reappears as kinetic energy of ejection.

The mass of the neutron has been evaluated by Chadwick, Feather and Bretscher by the following method. Neutrons are produced when deuterium is irradiated with the $2\cdot62\times10^6$ e.v. gamma radiation from Th C". The reaction is $^2_1\mathrm{D}+h\nu=^1_1\mathrm{H}+^1_0n$. The energies of the protons thus formed are determined by measuring their ranges or the amount of ionisation they produce. Since the neutron and proton masses are nearly equal it can be concluded that the neutron energies are about equal to the proton energies. Since the masses of $^2_1\mathrm{D}$ and $^1_1\mathrm{H}$ are known, and also the energies of the particles and quantum, solving this gives the mass of the neutron as $1\cdot00895$. This is slightly greater than the proton mass, $1\cdot00778$.

Bohr's theory of nuclear disintegration

A number of theoretical difficulties in connection with nuclear disintegration have recently been removed by Bohr, who has postulated a theory explaining the mechanism of disintegration. It had previously been assumed that in a disintegration process the projectile simply knocked a particle out of the nucleus leaving behind a changed nucleus. This apparently simple mechanism was difficult to apply in detail.

Bohr first assumed that the close nuclear packing makes it necessary to treat the nucleus as a whole, somewhat in fact as a drop of water is to be considered as an entity. When a projectile strikes a nucleus it is first captured, and since it is very energetic the effective temperature of the nucleus rises, possibly by several million degrees. This newly created intermediate nucleus has a great excess of thermal energy, and after a small but finite time the energy is distributed amongst all the nuclear particles. Calculations show that because of the high density of matter in the nucleus, the thermal conductivity is very high, thus the thermal energy is very rapidly distributed. If sufficient energy has been introduced, one of the particles already with large kinetic energy may receive a sufficient addition to enable it to overcome the potential barrier. This particle escapes and we have a disintegration.

Bohr has suggested the following model to illustrate the mechanism. Suppose there are a number of balls in a saucer, all moving, but none of which has sufficient velocity to carry it over the edge of the saucer. This represents a nucleus, the balls being the constituent particles, the height of the saucer wall being a model for the potential barrier within the nucleus. Suppose, now, a ball is shot into the saucer from outside (representing the projectile). If there are sufficient balls in the saucer the projectile will not be able to pass straight through but will be stopped by striking some of the balls already there. The energy of the projectile will rapidly be distributed amongst the rest of the balls, the velocities of which therefore increase. One of the balls near the edge of the saucer may easily quickly acquire sufficient energy to enable it to leave altogether. This represents the mechanism of disintegration by capture.

In the light of this theory the reaction for the production of neutrons by a-particles previously described should be modified as follows. First the boron nucleus captures the a-particle to form an isotope of nitrogen thus

$$^{11}_{5}B + ^{4}_{2}He = ^{15}_{7}N.$$

The nitrogen isotope, the intermediate nucleus, is unstable because of excess energy, and after a small but finite time disintegrates into ordinary nitrogen and a neutron thus

$$_{7}^{15}N = _{7}^{14}N + _{0}^{1}n.$$

The end of the reaction is the same as if there had not been an unstable intermediate nucleus.

Nuclear transmutations using accelerated particles

An important practical application of Gamow's theory is the way in which it pointed to the possibility of disintegrating nuclei by the aid of artificially accelerated particles. a-particles used for disintegration purposes have energies up to 8×106 e.v. In spite of this only one particle in a hundred From the thousand is able to produce a transformation. classical point of view at least several millions of volts are necessary before a disintegration of a medium-light nucleus can be produced, but Gamow's theory showed that there is finite probability that a much less energetic particle can penetrate a nucleus and as a result of capture produce disintegration. The probability diminishes with the energy, so that if an attempt is made to disintegrate nuclei with particles of relatively small energy, very large numbers will be required in order to produce an appreciable number of disintegrations.

This idea of using very large numbers of less energetic particles was first successfully put into practice in 1932 by Cockcroft and Walton, who were able to induce transformations in lithium and boron by bombarding the latter with a stream of protons (hydrogen ions) artificially accelerated with only

300,000 volts. In these disintegrations a-particles were ejected. Later experiments showed that, providing sufficient protons were available, disintegrations could be induced with potentials as low as 20,000 volts. This is a very interesting confirmation of the wave mechanics conception of nuclear structure.

The production of high voltages

Since high-voltage acceleration of particles plays an important part in disintegration investigations, we shall consider here briefly the methods which have been developed for this purpose. There are seven main methods for producing high voltages, all of which involve difficult problems in electrical engineering. The different methods, in brief, are as follows:—

1. The impulse generator

A system of condensers is charged in parallel and then discharged in series. The arrangement is relatively cheap and by it about 3×10^6 volts can be reached. However, the impulse only lasts 10^{-5} second, and as there must be 5 seconds' rest period between each discharge and the next the available energy is not great.

2. Condenser-rectifier voltage multiplier

Condensers and rectifying valves are arranged so that a transformer voltage can be multiplied then rectified. Some 2×10^6 volts can be obtained.

3. Electrostatic generators

The Van de Graaf generator consists of rapidly moving endless belts on to which charge is directed by corona points operating at 20,000 volts. The charge is carried to a large sphere and accumulates there until a potential of 5×10^6 volts builds up. The spheres used are 15 ft. in diameter and 22 ft. from the ground. The main problem lies in insulation. The apparatus must be contained in a large, high building otherwise discharges to the walls may occur. By the use of compressed air in the apparatus room still higher voltages can be achieved. It has been observed that gaseous $\mathrm{CCl}_2\mathrm{F}_2$ has remarkable insulating properties, being even superior to transformer oil at high pressures. The apparatus can be surrounded

by this gas to reduce the chances of a short circuit. The current available is decided by the speed of the belts. The machines already in use can produce 1 milliamp at 5×10^6 volts (an energy of 5 kilowatts).

4. High-frequency resonance transformers

A modified Tesla high-frequency discharge from a 200-kilowatt oscillator is passed through a single turn of copper tubing. This acts as the primary of a transformer. The secondary has fifteen turns, and both the coils are tuned to resonance. Oscillations up to 800,000 volts can be built up in the secondary coil.

5. Cascade transformers

Four transformers, each capable of producing 250,000 volts, are arranged in series. By this means 1×10^6 volts can readily be built up. The cost of the transformers is high.

6. Linear resonance acceleration

Charged particles are accelerated down the axis of a series of cylindrical electrodes which are alternately connected to the terminals of a high-frequency high voltage oscillator. The frequency and electrode lengths are adjusted so that as the charged particle emerges from each electrode the field has swung round enabling the next electrode to give the particle another accelerating impulse. By applying about 80 kilovolts an effective voltage of 2.85×10^6 has been attained.

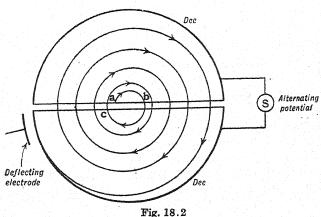
7. Magnetic resonance

This, the most important method, is that employed in the cyclotron, the theory of which will shortly be discussed. With all the other direct methods involving the actual production of true high voltages, it is considered that insulation difficulties set 6×10^6 as the upper limit. With the cyclotron the particles acquire the equivalent energy which would be produced by high voltage, without actually using a high potential. The equivalent of 19×10^6 volts has already been reached and with the new cyclotrons under construction still higher values are expected. The cost of the machine and operation costs are both high, and the working is difficult.

The cyclotron

A notable development in the method of disintegrating by accelerated particles has been brought about by Lawrence, who in 1932 devised a highly ingenious apparatus, now named the *cyclotron*, by means of which projectiles with energies up to 19×10^6 e. volts can be produced in very large numbers.

In the cyclotron the ions are energised by multiple acceleration, according to the principle illustrated in Fig. 18.2. The principal portion of the apparatus consists of a flat, circular, metal box cut into halves, the two parts being separated by a gap. Since they have the shape of the letter D they are called the "dees." An alternating potential of high frequency,



but in general not exceeding 100,000 volts peak value, is applied between the dees. These are in an evacuated chamber which is maintained in a uniform magnetic field of about 15,000 gauss. To do this magnets which may be 100 tons in weight are required and the consumption of electrical power is very high. Gaseous ions are generated near the centre of the dees either thermionically or else by special capillary ion sources of high intensity. The ions may be either protons, or deuterons or ionised helium atoms, etc.

The ions are accelerated by the potential difference between the dees, but since the latter are in a magnetic field, with the lines of force perpendicular to the dee plane, a small semicircular path is described. From what follows it will be seen that the time taken to describe a semicircle is independent both of the velocity of the ions and of the radius of curvature of the path.

The radius of curvature of the path R is given by

$HEV = MV^2/R$

where E, M, and V are the charge, mass, and velocity of the ion and H the strength of the magnetic field.

Thus

R = MV/HE.

Clearly the time T taken to traverse a semicircle is

 $\pi R/V = \pi M/HE = T$.

This is therefore independent both of R and V and is the same for all particles with the same mass and charge. The frequency of the applied alternating potential and the value of H can be chosen so that a particle starting at a just reaches b at the time when the potential has swung through half a cycle. The particle is therefore accelerated across the gap into the opposite dee. Since this act of acceleration increases the velocity of the ion it travels in a semicircle with larger radius and reaches the point c. But the time taken is independent of both radius and velocity so that c is reached exactly at the moment when the changing field can accelerate the particle once more across the gap. By this means several hundred separate accelerating impulses can be given to each particle before the outer rim of the dees is reached. At the outer circumference a supplementary electrode is placed which deflects the accelerated ions on to a target containing the material to be studied.

By substituting in the expression for T the value of E/M for a proton we find $T = (6.57 \times 10^{-4})/H$ sec. With a magnetic field H equal to 15,000 gauss the time taken is about 4.4×10^{-8} sec. The reciprocal of twice this, about 1.1×10^{7} , gives the frequency in cycles per second of the alternating potential which must be applied for the phase to change over exactly in time with the arrival of the particles at the gap. Such a frequency is that of a radio wavelength of about 26 metres.

For deuterons, twice this wavelength is required.

The final energy attainable depends on the magnetic field strength and upon the radius of the final path. This explains the need for powerful magnets giving a large field uniform

over a considerable area. With a radius of, say, 38 cms., and a field of 15,000 gauss, protons acquire a velocity of emergence of 5.4 × 109 cms./sec. which is equivalent to accelerating an electron by 15×106 volts. It would seem at first that there would be no limit to the increase of effective voltage attainable, but this is not so. The speeds of the particles are becoming so great that the relativistic increase in mass is significant. This leads to a steady increase in time in successive paths within the dees and the particles get out of step and are lost from the beam. There is thus at present an upper limit. The largest cyclotron yet constructed has a pole face diameter of 145 cms. and with this the equal to 19×106 electron volts has been reached. Still larger cyclotrons are projected and still higher energies are expected. The currents of these high-energy ions are extremely intense compared even with the strongest available radioactive sources. Thus some 90 microamps of deuterons of 16×106 electron volts can be produced and by irradiating beryllium with these a yield of 2×1012 neutrons per second results. It will be remembered 1 gram of radium gives 3.4×10^{10} a-particles per second, each of considerably less energy than the ions in the cyclotron So energetic are the beams that the ions have a range in air of over half a metre, forming a striking luminous beam when allowed to emerge through a thin window. The production of neutron radiation is so intense that special precautions to protect operators must be taken.

Artificial radioactivity

Up to 1933 radioactivity had always been considered to be a type of energy manifestation entirely beyond control, but in that year Mme. and M. Curie-Joliot succeeded in producing radioactivity by artificial means. In their experiments aluminium was bombarded with α-particles, a nuclear transformation taking place with an accompanying emission of neutrons. The reaction is

$$_{13}^{27}$$
Al $+_{2}^{4}$ He $=_{15}^{30}$ P $+_{0}^{1}$ n.

The residual atom, $^{30}_{15}$ P, is an isotope of phosphorus which does not normally exist. This newly created isotope is unstable and

radioactive, breaking down into a stable isotope of silicon by the spontaneous emission of a positive electron thus

$$_{15}^{30}$$
P= $_{14}^{30}$ Si+ e^+ .

The "radio-phosphorus," as it is called, has a normal radioactive decay curve with a period of 195 seconds and can be

proved to be phosphorus by chemical means.

This important experiment first proved that new radioelements could be created by bombarding *stable* atoms with high-speed particles. A rapid extension of the work followed, other projectiles such as accelerated protons, accelerated deuterons (the nuclei of deuterium), neutrons, etc., being employed. Neutrons are particularly effective in inducing nuclear transformations for, being uncharged, they penetrate nuclei very easily.

New radio-elements have now been created in large numbers and more are being discovered. Some of these new elements have positive electron emission, and others negative electron β -emission. Some of the ejected electrons have energies as high as 10^7 e.v. Exactly as in the case of natural β -ray activity the emission is in the form of a continuum with a sharp upper limit. The transformation periods observed vary over a wide range.

Nuclear chemistry

The production of artificial radioactivity only differs from the earlier transmutations already discussed in that the resultant nucleus is in one case stable and in the other case is unstable and breaks up spontaneously, due probably to a surplus of energy. Similar principles therefore govern both types of transmutation. In general a broad distinction can be made between two types of disintegration, namely, those in which the transformations are produced by the capture of the projectile and those in which they are produced without capture. The disintegration without capture is caused by the violent disturbance induced in a nucleus by the close passage of a charged projectile. In both types of disintegration there can be either emission of γ -radiation, or of particles, or of both.

The particular reactions which take place depend upon both the nature of the projectile and the nucleus which is struck. A given projectile is capable of producing usually quite a number of different reactions with a given nucleus, a fact which can be explained by Bohr's theory of disintegration, by what is effectively "evaporation" of a particle. We shall consider first transformations that can be induced with accelerated protons, since these are the simplest projectiles available.

A typical reaction with accelerated protons is obtained with lithium. Consider for the moment the heavier isotope only.

$$_{3}^{7}\text{Li} + _{1}^{1}\text{H} = 2_{2}^{4}\text{He}.$$

This reaction illustrates a case of *simple capture* with the formation of stable nuclei. The proton has entered the lithium nucleus and remained, producing two a-particles. A more complex example of capture is afforded by the impact of a proton on boron. The first possible reaction is simple capture, as above, leading to reaction,

$$^{11}_{5}B + ^{1}_{1}H = ^{12}_{6}C.$$

A stable isotope of carbon has therefore resulted. But other reactions are also possible and the following takes place:

$$^{11}_{5}B + ^{1}_{1}H = ^{8}_{4}Be + ^{4}_{2}He.$$

However, the ${}_{4}^{8}$ Be nucleus is *unstable* and tends to dissociate into 2 ${}_{0}^{4}$ He so that the above reaction ultimately becomes

$$_{5}^{11}B + _{1}^{1}H = 3_{2}^{4}He.$$

We shall now consider transmutations induced by neutrons of which there are very many examples. The following is an illustration of the number of possible reactions that can occur in a *simple* case. The resulting reactions induced by the bombardment of nitrogen with neutrons are:

(1)
$${}^{14}_{7}N + {}^{1}_{0}n = {}^{11}_{5}B + {}^{4}_{2}He$$
;

(2)
$${}^{14}_{7}N + {}^{1}_{0}n = {}^{14}_{6}C + {}^{1}_{1}H$$
;

(3)
$${}^{14}_{7}N + {}^{1}_{0}n = {}^{7}_{3}Li + 2{}^{4}_{2}He$$
.

In the last reaction three particles are produced. These

have been observed as a triple forked track in an expansion chamber.

The reactions produced by particles more complex than protons and neutrons are in no way different. The example below illustrates what takes place when the heavier isotope of lithium is bombarded with deuterons. There are again three possible reactions:

- (1) ${}_{3}^{7}\text{Li} + {}_{1}^{2}\text{D} = {}_{4}^{8}\text{Be} + {}_{0}^{1}n$;
- (2) ${}_{3}^{7}\text{Li} + {}_{1}^{2}\text{D} = 2{}_{2}^{4}\text{He} + {}_{0}^{1}n$;
- (3) ${}_{3}^{7}\text{Li} + {}_{1}^{2}\text{D} = {}_{3}^{8}\text{Li} + {}_{1}^{1}\text{H}$.

A reaction of particular interest takes place when accelerated deuterons are projected into deuterium gas. Two possibilities can occur, namely,

- (1) ${}_{1}^{2}D + {}_{1}^{2}D = {}_{1}^{3}H + {}_{1}^{1}H$;
- (2) ${}_{1}^{2}D + {}_{1}^{2}D = {}_{2}^{3}He + {}_{0}^{1}n.$

In the first case a new hydrogen isotope of mass three has been created, and in the second case a new helium isotope of mass three.

Finally, we shall consider two cases of disintegration, using radioactive a-particles as the projectiles, namely, the reactions produced by bombarding magnesium. • Consider only the two magnesium isotopes of masses 24 and 25. The reactions are:

$$^{24}_{12}\text{Mg} + ^{4}_{2}\text{He} = ^{27}_{14}\text{Si} + ^{1}_{0}n$$

 $^{25}_{12}\text{Mg} + ^{4}_{2}\text{He} = ^{28}_{13}\text{Al} + ^{1}_{1}\text{H}.$

It so happens that both the newly created ²⁷₁₄Si and ²⁸₁₃Al are unstable and radioactive, one emitting a positron the other an electron. The two decay processes are respectively

$$_{14}^{27}$$
Si= $_{13}^{27}$ Al+ e^+
 $_{13}^{28}$ Al= $_{14}^{28}$ Si+ e^-

It has only been possible to select a few examples of the effects produced by different projectiles. A whole new field of nuclear chemistry is rapidly being built up and it is quite

certain that much light will be thrown upon the structures of complex nuclei by the study of these reactions. There is a distinct possibility also of obtaining sufficient synthetic radioactive material to replace that naturally occurring. Many hundreds of new types of nuclei have up to date been created, most of them unstable and radioactive.

Nuclear isomerism

Chemical isomerism, frequently found in organic compounds, is due to the spatial rearrangement of a group of atoms forming a molecule. An analogous "nuclear isomerism" has now been found in a small group of artificially produced nuclei. These nuclear isomers have the same mass and charge but exhibit different radioactive properties. The suggested explanation is that the lowest excited state of a nucleus has a spin differing by many units from the still lower unexcited ground state. Theory shows that certain selection rules operate with the result that the first excited state is metastable. The nucleus can remain in this metastable state for a sufficient length of time for it to be observed as a different atom. Chemical separation of the isomers can sometimes be achieved. At least seventeen pairs of isomers have so far been detected.

The fission of uranium nuclei

A distinctly new type of disintegration has recently been observed when uranium and thorium nuclei are bombarded with neutrons. When these heavy atoms are irradiated new radioactive series are set up, at first thought to be "transuranic," that is, to involve species with atomic numbers greater than 92. Hahn and Strassmann showed, however, that the bombardment of uranium by neutrons leads to the production of barium and krypton nuclei by a process of "nuclear fission." A neutron is captured by the already complex and relatively unstable uranium nucleus and forms a very unstable nucleus of greater mass. This has a mean life less than 5×10^{-13} second. The close packing of the particles gives the nucleus properties analogous to those of a liquid drop, in accordance with Bohr's theory. The high excitation due to the neutron

capture causes the drop to split up into two large sub-particles which repel each other violently. Owing to the packing fraction differences there is a net gain in kinetic energy of the order of 200×10^6 e.v. per disintegration. This is of very great interest and may prove to be of importance, since it is clearly a case of tapping the enormous energy sources latent in nuclei. There is some evidence indicating that several neutrons may be "evaporated" from the fission products immediately after the fission has taken place.

The nuclear origin of stellar energy

In stellar interiors the temperatures are very high, so great indeed that the thermal velocities of atoms are sufficient to lead to nuclear disintegration when collisions occur. It can be shown quantitatively that these "thermo-nuclear reactions" account for the tremendous emission of energy from stars. Stars can be roughly divided off into two groups brighter and fainter than our sun. As shown spectroscopically both types have a high hydrogen content. In the hot interiors all atoms are stripped of electrons by thermal ionisation. The calculated temperatures are ultimately of the order of 20,000,000° C., which considerably exceed those needed for nuclear disintegration. The high concentration of energetic protons leads to nuclear disintegrations with considerable emission of energy. Approximately correct numerical results are obtained by assuming that a "carbon-nitrogen disintegration cycle" takes place in the brighter stars and a proton-proton interaction in the fainter.

In the bright stars the carbon acts as a catalyst converting hydrogen into helium according to the following chain of reactions, with emission of positrons and γ -radiation:

$$^{12}_{6}\text{C} + ^{1}_{1}\text{H} \rightarrow ^{13}_{7}\text{N} + \gamma \text{ radiation}$$

$$^{13}_{7}\text{N} \rightarrow ^{13}_{6}\text{C} + e^{+}$$

$$^{13}_{6}\text{C} + ^{1}_{1}\text{H} \rightarrow ^{14}_{7}\text{N} + \gamma \text{-radiation}$$

$$^{14}_{7}\text{N} + ^{1}_{1}\text{H} \rightarrow ^{15}_{8}\text{O} + \gamma \text{-radiation}$$

$$^{15}_{8}\text{O} \rightarrow ^{15}_{7}\text{N} + e^{+}$$

$$^{15}_{7}\text{N} + ^{1}_{1}\text{H} \rightarrow ^{12}_{6}\text{C} + ^{4}_{2}\text{He.}$$

The ¹²₆C remains practically unaffected enabling the cycle to continue for great time periods with consumption of protons only.

In the fainter stars it is probable that the main reaction is simply ${}_{1}^{1}H + {}_{1}^{1}H \rightarrow {}_{1}^{2}H + e^{+}$. Intermediate cases, such as that of our sun, probably involve both reactions to about the same extent.

It is to be noted that these modern views on stellar energy emission and stellar evolution imply that any heavier elements in a star, even such as Li, Be, B, etc., are destroyed by proton collision, with resulting a-particle emission. Such atoms, on this view therefore, existed before the star as such was formed. The evolution of a star is now considered to be as follows. A star begins by the condensation of a large amount of matter at a point in space. Owing to gravitational attraction this cluster gradually contracts, the gravitational potential energy thus set free being ultimately liberated as radiation. When contraction proceeds the temperature rises until that of the core is about 200,000° C. At this point the thermal energy suffices to enable interactions to take place between protons and any deuterons present. Calculation shows that after some 106 years all the deuterons will be exhausted, gravitational contraction then being resumed for a period of 103-105 years.

As a result of the second contraction the core temperature rises to about 5×10^6 °C., the thermal velocities of the protons being then great enough to permit of interactions with deuterium, lithium, beryllium, and boron. This stage of the star's life is described as the "giant stage." When all the above four types of nuclei have been consumed further contraction sets in, the temperature rising to 20×10^6 °C. When this is reached the carbon cycle begins, supplying energy for the main radiating life of the star. The hydrogen is steadily converted into helium and when the protons are largely exhausted another contraction begins, this time rapidly. The energy so liberated soon exceeds the thermo-nuclear energy. The centrifugal forces increase and the star may then break up.

This theory of stellar evolution does not account for the presence of the original cosmic matter forming the star, nor does it give any clue to the mode of formation of complex heavy atomic nuclei.

Nuclear spin

Important contributions to nuclear theory have resulted from spectroscopic studies of nuclear spins. Nuclei can be divided off into four distinct groups according to the nuclear mass and charge. About half the isotopes in the Periodic Table have even nuclear mass and even nuclear charge, about a quarter have odd mass and even charge, and another approximate quarter have odd mass and odd charge. The fourth group consists of four species only in which the mass is even and the charge odd.

The first important conclusion derived from hyperfine structure observations is that practically all the nuclear spins determined with certainty are half integral and are associated with odd atomic weights. (The small fourth group is an exception to this rule and will be discussed separately.) A second conclusion is that light nuclei with masses integral multiples of four have almost certainly zero spin. This has been proved for the lighter elements, helium, carbon, oxygen, and sulphur. This fact is important, for these particular elements lie on the lower spur of the mass defect curve. Spin data and mass defect data seem to imply the existence of the a-particle as a nuclear sub-unit. Since the a-particle has a zero nuclear spin, nuclei built up only from them will also have zero spin.

It is very likely that nuclei consist only of neutrons, protrons, and possibly a-particles formed from groupings of the latter. Clearly a nucleus with an odd mass number has an odd particle left over in its constitution. If the nuclear charge is odd this odd particle must be a proton, but if the nuclear charge is even it must be a neutron. One can therefore distinguish more precisely the two groups of nuclei with odd atomic weight, for in one there is an odd proton and in the other an odd neutron.

Observed nuclear spins vary from 1/2 to 9/2 in spite of the fact that a nucleus can contain, when heavy, more than 200 particles. Theory and experiment agree that a proton and a neutron both have inherent spins of 1/2. (The magnetic moment of such a particle will be considerably less than that of an electron because of the greater mass.) Landé has therefore proposed a theory to explain the origin of the nuclear

spins. In this theory Landé considers that the nuclear spin is to be attributed to the single odd particle, neutron or proton. This particle has a spin of 1/2 and is also assumed to have an orbital angular momentum which may have a quantum number 0, 1, 2, 3, etc., exactly analogous to the l value of the electron in its orbit. The spin of the odd particle, together with its angular momentum, combine to form a resultant which is the nuclear spin. This theory of Landé succeeds in explaining why only small nuclear spins are observed. For example, the nuclear spin for arsenic has been stated to be 3/2 and this could arise from the addition of the spin of the odd proton either in parallel with an orbital momentum of 1 or anti-parallel with an orbital momentum 2.

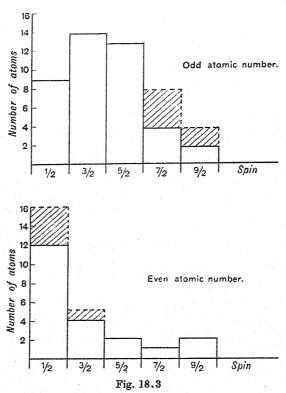
The theory is able to show that the nuclear magnetic moment will be different in the two cases. The calculations made on this basis for a number of nuclear magnetic moments are in fair agreement with those observed. Even when the hyperfine structures have been accurately measured it is difficult to derive therefrom exact values for nuclear magnetic moments. It is necessary in some cases to assume that three neutrons can contribute to the final resultant spin.

The distribution of nuclear spins

It has been shown by Tolansky that the numerical distribution of nuclear spins differs in a striking manner for atoms with odd nuclear proton and those with odd nuclear neutron. The two forms of distribution are shown in Fig. 18.3. The shaded areas represent atoms for which a doubtful, but probable, value of the spin has been deduced. Smaller spins are favoured in the atoms with even atomic number (odd nuclear neutron). The general form of the distributions can be accounted for by the Landé theory of nuclear spins. These curves show that the nature of the odd particle decides the spin and this is precisely what the Landé theory predicts.

The distribution of spins in the small fourth group of atoms, those with even mass and odd charge, is of special interest. The stable atoms in this group are ${}_{1}^{2}D$, ${}_{3}^{6}Li$, ${}_{5}^{10}B$, ${}_{7}^{14}N$. The nuclear spins are known for these. They are each integral and

equal to unity. The deuteron is structurally a combination of one proton and one neutron and as each of these constituents has a spin of 1/2 it can be considered that in the deuteron the spins of the individual components add up in parallel to give a resultant spin of 1. The lithium isotope of atomic weight 6 probably contains an α -particle together with a neutron-proton pair in the nucleus. The spin of the α -particle is zero,



since the nuclear spin of helium is zero. The remaining proton-neutron pair will be expected to have the same residual spin as in the case of the deuteron, namely, unity. This is indeed the case. By a similar argument, the $^{10}_{5}$ B nucleus will contain two a-particles and a proton-neutron pair, and the $^{14}_{7}$ N nucleus three a-particles and a proton-neutron pair. This simple conception therefore predicts that the spins of all the four

members of the group will be identical and equal to unity, a prediction which is in complete agreement with observation.

It is to be noted that the Landé theory implies the existence of extremely small orbits within the very small radius of the nucleus, which is of the order of 10^{-13} cms.

Atomic particles, nuclear particles and their transitions

An analogy can now be drawn between the properties of atoms and nuclei. This is shown in the following table, suggested by Heisenberg:—

	Outer atom	Nucleus proton neutron α-particle	
Constituent particles which can be emitted.	electron		
Emission during transitions.	photon	y-ray (photon) electron positron neutrino meson { positive } negative }	
Corresponding field.	electromagnetic	electromagnetic De Broglie waves	
Type of force.	Coulomb field	Coulomb field exchange force	

It is evident that nuclear phenomena are much more complex than atomic phenomena. More particles may yet be discovered. There is, for example, always the possibility of a negative proton being found.

REFERENCE

[&]quot;Report on Progress in Physics." Vol. 4 (1940). Physical Soc. of Lond.

[&]quot;Introduction to Nuclear Physics." N. Feather. (1936.)

CHAPTER 19

THE RELATIVITY THEORY

The Michelson-Morley experiment

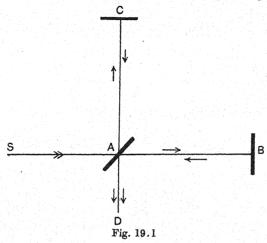
The Theory of Relativity, propounded by Einstein in 1905, and the quantum theory of Planck, have completely revolutionised modern scientific and philosophic thought. In the present chapter the restricted or so-called *Special Theory of Relativity* will be discussed. This is based upon a celebrated optical experiment carried out in 1897 by Michelson and Morley in an attempt to measure the drift of the ether past

the earth, as it moves in its orbit round the sun.

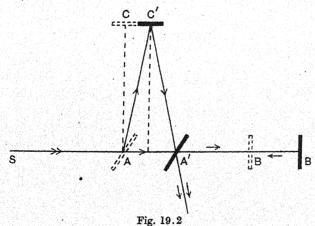
Physical theory, particularly that associated with the propagation of electromagnetic waves, at that time postulated the existence of a universal all-pervading ether filling space. The earth, in its orbital motion, passes through this ether with a high velocity. If, therefore, a beam of light is sent from a source to a receiver on the earth in the same direction as the earth's motion, it should take longer to complete the journey than if sent in the reverse direction, against the earth's motion. This is because, on this view, the light travels in the ether with a fixed velocity, and the receiver is moving forward in front of the advancing wave when emission is in the same direction as the earth's motion, and vice versa. The aim of the experiment was the measurement of this time difference, from which the relative velocity of drift between the ether and the earth can be calculated.

To make the observations a Michelson interferometer was set up as in Fig. 19.1. A beam of monochromatic light SA was partly reflected and partly transmitted by the half-silvered mirror at A. The divided beams fell normally on the mirrors C and B and returned to A where they united to form interference bands, observable by an eye placed at D. Let it be supposed that the paths AC and AB are of equal length d.

Suppose the whole apparatus moves through the ether with velocity v along SB. (v is in this case the velocity of the earth in the direction SB.) Let c be the velocity of light through



the ether. Owing to the velocity of the system the optical path becomes that shown in Fig. 19.2. The beam reflected from A strikes the upper mirror at an angle other than the normal, since the mirror has moved from C to C' in the time



taken for the beam to reach it. Let T be the time required for the beam to pass from A to B and T₁, that needed to pass back again to A which has now moved to A'. We have

$$T = d/(c-v)$$

$$T_1 = d/(c+v).$$

The total time for the journey and back is

$$T+T_1 = \{d/(c+v)\} + \{d/(c-v)\}$$

$$= 2dc/(c^2-v^2).$$

Since the distance travelled by the light is the time taken times the velocity of light, the length of the path is

$$\begin{array}{l} 2dc^2/(c^2-v^2) \\ = 2d/(1-v^2/c^2) \\ = 2d(1+v^2/c^2) \end{array}$$

to a first approximation.

The length of the path AC' can easily be derived from Fig. 19.2 for AC=d and also C'C/AC'=v/c, hence we have

$$(AC')^2 = AC^2 + (C'C)^2$$

 $(AC')^2 \left(1 - \frac{v^2}{c^2}\right) = AC^2$
 $AC' = d/\sqrt{1 - \frac{v^2}{c^2}}$

or to a first approximation

$$AC' = d(1 + v^2/2c^2).$$

Since AC'=C'A' the total path AC'A'equals $2d(1+v^2/2c^2)$.

There is therefore a difference dv^2/c^2 in length between the path from A to B and back and the path from A to C and back. If the earth were to stop suddenly the interference fringes would be displaced, because the path difference would then vanish. If the whole apparatus is turned through 90° the arm AC now becomes optically longer than AB by dv^2/c^2 instead of being the shorter. The effect of rotating the apparatus is to produce a path difference equal to $2dv^2/c^2$. From the known motion of the earth the ratio v/c is approximately 10^{-4} . If d is 10 metres the optical path difference is 2×10^{-5} cms. This is 2/5 of a wavelength of sodium light, an amount very easy to observe since a displacement of only one-hundredth of a wavelength can just be detected.

The result of the Michelson-Morley experiment

To reduce the dimensions of the apparatus in order to maintain the temperature constant and avoid vibrational effects, the arms AC and AB were kept small, but d was increased to 10 metres by a system of mirrors reflecting the beams to and

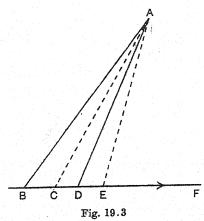
fro before returning to the half-silvered mirror. The whole optical system was mounted upon a massive base which floated on a pool of mercury. The apparatus could be kept in slow rotation, at a rate of ten complete rotations an hour.

The result found by Michelson and Morley was surprising. There was no observable displacement of the fringes, suggesting that the relative velocity between the earth and the ether is zero, or certainly less than a fortieth of that expected. There was, however, just the possibility that at the time of the experiment the real velocity of the earth in space was very small, because of the motion of the solar system as a whole. At a given time of the year it would be possible for the orbital and solar motions to be in opposite senses and practically cancel. To eliminate this possibility, observations were made at different times of the year, when the directions of the earth's orbital motion are different. In no case was there an observable ether drift.

Aberration of light and ether drift

When the result of the Michelson and Morley experiment is closely considered it is seen to reveal a remarkable discrepancy with earlier experiments made to test ether drift. One of the most striking of these is Airy's experiment on the aberration of The fact that aberration occurs at all appears to show that the earth is in motion through the ether: aberration is the apparent angular change in direction of a star due to the velocity of the earth. Airy reasoned as follows. Supposing the ether to be at rest, then, if a star is viewed with a telescope whose barrel is filled with water of refractive index μ , a greater aberration will be produced because, although the velocity of the earth is not changed, the velocity of the light down the telescope axis is reduced by the water. Observation shows, however, that the aberration is not affected by the introduction of the water medium. It was shown by Fresnel that the only "classical" explanation for this is that the ether drifts with the water in the telescope with a velocity $v(1-1/\mu^2)$ where v is the velocity of the earth in its orbit. The following derivation will prove this.

In Fig. 19.3 let AD be the *true* direction of incidence of light from a star falling upon the object-glass of the telescope A. Owing to the velocity of the earth in the direction BF, AB is the apparent direction. The angle BAD is the aberration when the telescope is filled with air. If the telescope tube is now filled with water, the rays incident at A are refracted by the water to the direction AC, AD being the direction of incidence and AB the normal. From this we have $\sin \text{BAD} = \mu \sin \text{BAC}$, or as the angles are small, $\text{BD} = \mu \text{BC}$. Since the velocity of light in water is $1/\mu$ that in air, the time taken to traverse the water-filled telescope must be μ times that taken to traverse the same telescope when filled with air.



The eyepiece will therefore have arrived at E, where $BE = \mu BD$, when the ray is expected to arrive at C.

Since observation shows that the aberration is the same as with an air-filled telescope, the ray must actually arrive at E and not at C. The only explanation is that the ether drifts the distance CE whilst the ray passes down from A. However, the water has moved through BE with the velocity v, hence the velocity of drift is

$$v' = CE \cdot v/BE$$

= $(BE - BC)v/BE$
= $v(1 - BC/BE)$
= $v(1 - 1/\mu^2)$.

For air the drift velocity is practically zero, since μ is nearly 1.

This formula, due to Fresnel, appeared to be completely verified in another manner as the result of an experiment by Fizeau in 1859 when the difference in the velocity of a beam of light sent with and against a moving stream of water was measured.

The Lorentz-FitzGerald contraction

The experiments just described establish the fact that when we are concerned with moving matter (water) there seems to be an ether drift, the velocity being that given by the Fresnel formula. On the other hand, the Michelson-Morley experiment appears to show that the ether surrounding the earth moves with

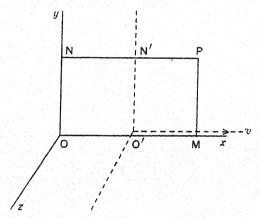


Fig. 19.4

the velocity of the earth itself. To remove this discrepancy Lorentz and FitzGerald independently proposed a remarkable hypothesis now known as the "Lorentz-FitzGerald contraction hypothesis." They suggested that the length of the path AB in Fig. 19.1, which lies in the direction of the earth's motion, contracts due to that motion by an amount $dv^2/2c^2$. If this happens the optical paths in the interferometer have the same equivalent lengths, leading to a null result for the Michelson-Morley experiment. The percentage contraction predicted is only about 5×10^{-7} per cent. The contraction in a length of 2 Kms. is only 10^{-3} cms.

The contraction theory is based upon electromagnetic conceptions which were considered justified because of the

electrical nature of matter. This will be made clear by a consideration of what is called the Lorentz transformation. Suppose we consider two observers, one O, at rest, and the other O' moving with a velocity v along the x axis. Let each describe the rectangular co-ordinates of the same point P (see Fig. 19.4). The motion is equivalent to two frames of reference (rectangular co-ordinates) in relative motion along the x axis. Let the figure represent the state of affairs at a time t measured by the clock of the observer O and t measured by the clock of O'. According to observer O the co-ordinates of P are x, y, z, and t and to O' they are x', y', z', and t'. Let O and O' coincide in space at time zero. Clearly, since NN' = vt and as PN' = x', it would appear obvious that

$$x' = x - vt$$

$$y' = y$$

$$z' = z$$

$$t' = t$$

One should be able to transform from the co-ordinates of the observer at rest to those of the observer in motion by the above simple relationships. This is described as the classical transformation for non-accelerated systems. When critically examined it is seen to contain the assumption that space and time are absolute measurements independent of each other and independent of the systems.

However, both experiment and theory show that if electromagnetic measurements stated with respect to fixed axes are transformed to axes moving with velocity v in the x direction, the form of the transformation rules changes. In order to maintain a principle of relativity it is necessary to be able to transform from axes at rest to axes in motion without changing the form of the equations of motion. Lorentz pointed out that the equation forms would be retained unchanged if instead of the classical transformation, one used

$$x' = \beta(x - vt)$$

$$y' = y$$

$$z' = z$$

$$t' = \beta(t - vx/c^2)$$

in which $\beta = 1/(1-v^2/c^2)^{\frac{1}{2}}$ where c is the velocity of light.

The differences between the classical and the Lorentz transformations are shown thus:

 $\begin{array}{ll} \textit{Mechanical transformation} & \textit{Lorentz transformation} \\ x' = x - vt & x' = \beta(x - vt) \\ t' = t & t' = \beta(t - vx/c^2). \end{array}$

According to the observer O the length he estimates for PN' is x-vt. This is less than the length x' which the observer O' attributes to the same distance since β is greater than unity. Each observer therefore ascribes a contraction to the measurement made by the other in the direction of relative motion. Furthermore, space and time are not independent entities but are related.

As $\beta=1/(1-v^2/c^2)^{\frac{1}{2}}$ when v is small compared with c this approximately equals $1+v^2/2c^2$. The contraction ascribed by each observer to the unit length measurements of the other is $v^2/2c^2$, exactly the amount required to explain the result of the Michelson-Morley experiment.

However, if the proposed contraction has a purely electromagnetic origin, Lord Rayleigh pointed out that transparent moving media should become doubly refracting owing to the stresses set up by the contraction. Although sought for, this effect has not been found. All the difficulties and anomalies have been removed by the Theory of Relativity propounded by Einstein in 1905. This shows that the electromagnetic contraction is but a special case of a general law applicable to all matter.

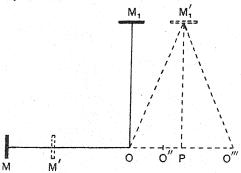
The Special Theory of Relativity

Perhaps the most remarkable feature about Einstein's theory is that it is based upon a single statement of fact arising out of the Michelson-Morley experiment. This experiment shows that the velocity of light measured by an observer is constant and does not depend upon his own velocity in the direction to and from the source emitting the light. This is the fundamental basic fact upon which the theory rests, namely, that the velocity measured for light is independent of the velocity of the observer. We shall now consider the implications of this.

Let MM_1 be two mirrors equidistant from an observer O (Fig. 19.5). Let M, M_1 and O be at rest with respect to each other and let O and O' have a relative velocity v in the direc-

tion MO. O sends out light signals simultaneously to M and M_1 and as $OM = OM_1$ the signals return simultaneously.

Let the observer O' know that the observer O receives the returned signals at the same instant. He is not aware of the fact that $OM = OM_1$, since the effective velocity of the light may be different in the two paths. Let him call the length in the direction of motion λ and that in the direction perpendicular



× O' Fig. 19.5

to the motion l. Let the full lines in the figure represent positions relative to O' initially, and dotted lines subsequent positions as seen by this observer. We shall follow out the deductions which O' would make on the assumption that the velocity of light, with respect to himself, is independent of direction.

According to the observer O' the length MM' equals vt_1 and OM' equals ct_1 if t_1 is the time taken for the light signal to go from O to the mirror M, which by this time has moved to M'.

OM' = OM - MM' i.e. $ct_1 = \lambda - vt_1$ giving $t_1 = \lambda/(c+v)$.

If t_2 is the time taken for the light to return from M' to the observer who is now at O''' then M'O'''= ct_2 . Since M'O''= λ and as O''O'''= vt_2

then $ct_2 = \lambda + vt_2$ i.e. $t_2 = \lambda/(c-v)$.

The total time taken for the light to go to the mirror M and come back is $T=t_1+t_2$.

Substituting gives

$$T = \lambda/(c+v) + \lambda/(c-v)$$

= $2c\lambda/(c^2-v^2)$.

So far we have only considered the signal sent in the direction of motion, we shall therefore now examine the behaviour of the signal sent in the perpendicular direction. OM_1' is equal to $O'''M_1'$. As the total time taken for the double journey is T, because the signals are received at the same time, the time to travel over OM_1' is T/2. The distance $OM_1'=cT/2$, and as O has travelled to P during this time (T/2) it follows that OP=vT/2. Since $OM_1'P$ is a right-angled triangle,

$$\begin{array}{ccc} & {\rm OM_1'^2{=}OP^2{+}M_1'P^2} \\ {\rm hence} & c^2{\rm T^2/4}{=}v^2{\rm T^2/4}{+}l^2 \\ {\rm giving} & {\rm T}{=}2l/(c^2{-}v^2)^{\frac{1}{2}}. \end{array}$$

The value just derived for T can now be equated to that already found for the other arm, i.e.

$$2\lambda c/(c^2-v^2) = 2l/(c^2-v^2)^{\frac{1}{2}}$$
 whence
$$\lambda = l(1-v^2/c^2)^{\frac{1}{2}}.$$

Thus, according to the observer O', the length l in the direction of motion contracts to λ where $\lambda = l(1-v^2/c^2)^{\frac{1}{4}}$. This is exactly the contraction required by the Lorentz transformation. It has been derived here mechanically and is not merely an electromagnetic property. It is easy now to derive the complete Lorentz transformation by referring back to Fig. 19.4. According to the observer O, PN'=x-vt where t is the time measured by him for O' to move the distance NN'. The observer O' calls the distance PN',x'. However, as just proved above, this appears to O to be equal to $x'(1-v^2/c^2)^{\frac{1}{2}}$ giving us

$$x'(1-v^2/c^2)^{\frac{1}{4}} = x-vt$$

i.e. $x' = \beta(x-vt)$ where $\beta = 1/(1-v^2/c^2)^{\frac{1}{4}}$.

This is the equation given by Lorentz for electromagnetic transformations.

The transformation for the time variable is derived from the same figure in the following simple manner. PN=PN'+NN',

and according to O' this is equal to x'+vt'. The length PN, which is x for the observer O, appears by the above reasoning to be $x(1-v^2/c^2)^{\frac{1}{2}}$ to the observer O' who therefore equates

$$\begin{array}{c} x(1-v^2/c^2)^{\frac{1}{2}}\!=\!x'\!+\!vt'\\ \text{which gives} & x\!=\!\beta(x'\!+\!vt'). \end{array}$$

Since, however, $x' = \beta(x-vt)$ we can substitute for x' and we find that

$$x = \beta \{\beta(x - vt) + vt'\}$$

= $\beta^2 x - \beta^2 vt + \beta vt'$.

Rearranging gives

$$\begin{array}{ccc} \beta vt' = \beta^2 vt + x(1-\beta^2) \\ 1 - \beta^2 = \{1 - 1/(1-v^2/c^2)\} \\ = (-v^2/c^2)/(1-v^2/c^2) \\ = -v^2\beta^2/c^2. \\ \text{Hence} & \beta vt' = \beta^2 vt - v^2\beta^2 x/c^2 \\ \text{giving} & t' = \beta(t - vx/c^2). \end{array}$$

By similar reasoning

$$t = \beta(t' + vx/c^2).$$

As lengths perpendicular to the direction of motion are not affected we finally obtain

$$\begin{array}{lll} x' = \beta(x - vt) & x = \beta(x' + vt') \\ y' = y & y = y' \\ z' = z & z = z' \\ t' = \beta(t - vx/c^2) & t = \beta(t' + vx'/c^2). \end{array}$$

This is the complete Lorentz transformation.

Physical significance of the transformation

We shall now derive a number of important physical conclusions by applying the above transformations. These hold for all matter since they have been derived in terms of dynamics only, without invoking any electromagnetic properties as postulated by Lorentz. It is clear that observers in relative motion ascribe a contraction to each other's lengths. This is of great importance and can explain many physical phenomena, such as the result of the Michelson-Morley experiment.

Observer's times as well as lengths are affected also. Con-

sider a clock on the system of O' at the origin, i.e. where x' is zero. From the above transformation $t=\beta t'$. However, t' is the time of the clock according to O' and t that according to O, and as β is greater than one, more time has elapsed according to O than according to O'. It follows that all movements on the reference frame of O' appear lethargic to O. The opinion is reciprocal, for a clock at the origin of the reference frame of O(x=zero) obeys the relationship $t'=\beta t$, leading to a similar conclusion. The clocks would appear infinitely slow if one of the observers were moving with the velocity of light.

Consider now a particle moving with constant velocity along the x axis. According to the observer O' the particle moves from the point x_1 at time t_1 to the point x_2 at time t_2 . He therefore ascribes to it a velocity

$$v_1' = (x_2' - x_1')/(t_2' - t_1').$$

In a similar manner the velocity of the same particle is estimated by the observer O to be

$$v_1 = (x_2 - x_1)/(t_2 - t_1).$$

Substituting from the transformation equations for x_1 , x_2 , t_1 , t_2 in the latter expression (remembering that v is the *relative* velocity of the two observers), leads to

$$v_1 = \frac{\beta(x_2' + vt_2') - \beta(x_1' + vt_1')}{\beta(t_2' + vx_2'/c^2) - \beta(t_1' + vx_1'/c^2)}$$

dividing through by β and then rearranging gives

$$v_1 = \frac{x_2' - x_1' + (t_2' - t_1')v}{(t_2' - t_1') + v(x_2' - x_1')/c^2}.$$

Dividing top and bottom by $t_2'-t_1'$ leads to

$$\begin{array}{l} , \;\; v_1 \! = \! \frac{(x_2' \! - \! x_1')/(t_2' \! - \! t_1') \! + \! v}{1 \! + \! \{v/c^2\}\!\{(x_2' \! - \! x_1')/(t_2' \! - \! t_1')\}} \\[1ex] v_1 \! = \! \frac{v \! + \! v_1'}{1 \! + \! v v_1'/c^2}. \end{array}$$

This is the law for the addition of velocities and may be compared with the classical or Newtonian law which states simply,

that if reference frames have a relative velocity of v and if the particle velocity estimated by O' is v_1 ', then that observed by O is just $v_1 = (v + v_1)$.

The velocity of light

The law for the addition of velocities leads to a very important conclusion relating to the maximum velocity attainable. Suppose that to an observer O' the particle has a velocity equal to c, the velocity of light. Since the observer O' is moving past O with a relative velocity v it is of interest to see what velocity O will now ascribe to the moving particle. This is obtained by putting $v_1'=c$ in the formula just derived. When this is done we get

$$v_1 = \frac{v+c}{1+vc/c^2}$$
$$= c.$$

A surprising fact therefore emerges. O also measures the velocity of the particle as the velocity of light. This is indeed in harmony with the fundamental assumption originally made, namely, that the velocity of light is independent of the velocity of the frame of reference. It is clear that the maximum velocity ever attainable is that of light, for the addition of the velocity v to a particle with the velocity of light c results in a final velocity c. This limit to the velocity attainable in nature is of fundamental importance.

The contraction in length in the direction of motion is also associated with the limit of attainable velocity. It has already been proved that a length l contracts by virtue of its velocity v to $l(1-v^2/c^2)^{\frac{1}{2}}$. If v has the value c, i.e. if the matter moves with the velocity of light its length becomes zero. This is clearly not possible, hence a particle can only approach but never attain the velocity of light. In the next section it will be shown that mass depends upon velocity. The mass of a particle becomes infinite when it acquires a velocity equal to that of light, another fact showing that this velocity is never actually attainable by matter. Therein lies the fundamental distinction between matter and quanta, since the latter do have the velocity of light.

The variation of mass with velocity

According to classical mechanics mass is constant under all conditions of velocity and acceleration. Relativity considerations show that this is not the case. Consider what takes place when two balls of equal mass (when compared together in the same reference frame) impact together at an angle. As shown

in Fig. 19.6 they move off after the impact with the horizontal components of the initial velocities unchanged and with the vertical components reversed.

Let the two observers O and O' have spheres of equal mass (compared at rest) and let the observer O' have a velocity v in the x direction. Each observer projects his sphere with (according to himself) the same speed V in the y direction so that they will meet half-way between the observers at the

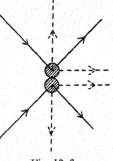
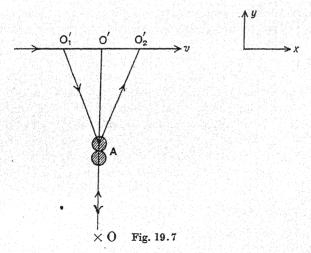


Fig. 19.6

point A (see Fig. 19.7) when O and O' are directly opposite each other on the y axis. To do this O' must project his



sphere from the point O_1 . As the x-component of velocity of his sphere is not affected by the impact it will return to the observer at the point O_2 such that O_1 $O' = O'O_2$. The y components

of velocity of the spheres are interchanged and, as both are V, each returns back to his own observer with the same velocity, namely, V.

O infers that the momentum in the y direction of the other sphere is equal to that of his own, since the latter returns with unchanged velocity. If then, in his reckoning, the mass and velocity of his sphere and that of the other observer are m, V, and m_1 , V_1 , respectively, he equates $mV = m_1V_1$. As the meeting point A is half-way between O and O', the distances OA and O'A are the same to both observers, being perpendicular to the direction of motion and, according to the transformation, not affected in length.

However, the *time* which O thinks is taken by the sphere from O' to reach A is β times longer than the time he considers to be taken by his own sphere. Hence to O the sphere of O' appears to move at a velocity slower than V, i.e. with a velocity V/β . Thus he equates $V_1 = V/\beta$. This leads to

$$m{
m V}\!=\!m_1{
m V}/eta, \ i.e. \qquad m_1\!=\!meta \ {
m or} \qquad m_1\!=\!m/(1\!-\!v^2/c^2)^{rac{1}{2}}.$$

This relationship shows that, according to O, the masses on the reference frame of O' increase with velocity, being infinite if the velocity of light is reached. This law relating mass and velocity has been directly confirmed by measurements of the masses of high-speed electrons. Very many observations made with both bound and free high-speed particles indirectly confirm this law which is now firmly established. The change of mass with velocity is very important in spectroscopic theory, for in elliptic orbits the velocities of the electrons are so great when near to the nucleus that the mass change with velocity affects the energies of the orbits considerably.

The four dimensional continuum

To conclude this survey of the Special Theory of Relativity, a brief indication of its further development to accelerated systems, called the General Relativity Theory, will be considered. The application of the Special Theory to the conception of a four-dimensional continuum made by Minkowski, is the inter-

mediate stage in the evolution of the gravitational theory of Einstein. When the equation of a line in Cartesian co-ordinates is transformed we arrive at the following.

The distance s between two points with co-ordinates $x_1y_1z_1$ and $x_2y_2z_2$ is given by

$$s^2 = (x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2$$
.

On applying the transformation equations we must expect to arrive at an equation of the same form in the co-ordinates of a second frame of reference, moving with respect to the first with velocity v. This is described by saying that the equation must be an invariant, i.e. unaltered in form by a change of axes. When the transformations already derived are applied to the above equation the result is

$$s^2 = \beta^2 \{ (x_2' - x_1') + v(t_2' - t_1') \}^2 + (y_2' - y_1')^2 + (z_2' - z_1')^2$$

which clearly depends upon v and is certainly not similar in form to the original equation.

It will be found by substitution that the expression

$$s^2\!=\!(x_2\!-\!x_1)^2\!+\!(y_2\!-\!y_1)^2\!+\!(z_2\!-\!z_1)^2\!-\!c^2(t_2\!-\!t_1)^2$$

does remain invariant when a transformation is made. For a line element this invariant form of equation is

$$ds^2 = dx^2 + dy^2 + dz^2 - c^2 dt^2$$
.

The units of length, etc., can be adjusted for mathematical simplicity so that the velocity of light c is unity (i.e. the unit of length is 3×10^{10} cms.). This simplifies the invariant slightly to

 $ds^2 = dx^2 + dy^2 + dz^2 - dt^2$.

Physically this means that a length is not described by the three dimensions dx, dy, dz, but in addition, the time dimension dt is forced upon us. The quantity ds is called a "point-event" and is not exactly a distance in space but is an interval in a four-dimensional continuum involving space and time inextricably bound to each other.

It will be noticed, from the first form of the invariant, that when ds=0 we have $c^2=(dx^2+dy^2+dz^2)/dt^2$, i.e. something moving with the velocity of light. This can only be a light

quantum (or a ray of light) hence the equation ds=0 is that for a ray of light.

In the four-dimensional continuum the equivalent of a straight line is a line making ds a minimum. This is termed a "geodesic." It is clear that the history of a particle can be represented by the integral of ds, such an integral being called a "world line." An ordinary "event" consists of the intersection of two world lines.

The General Theory of Relativity

A discussion of the General Theory of Relativity passes beyond the scope of this book. By extending the special theory to accelerating systems, Einstein has derived a law of gravity superior to that of Newton. The Newtonian inverse square law of gravitational attraction appears as a first approximation in the new law. Briefly speaking, Einstein assumes that the presence of matter distorts the curvature of the four-dimensional space. This ultimately leads to what is equivalent to gravitational attraction.

Three possible experimental confirmations of this abstruse

theory were predicted by Einstein, namely;

(1) It had long been known that the orbit of the planet Mercury exhibits a precession of 42" arc per century, a phenomenon which cannot be explained by the inverse square law of gravitation. Einstein found that the orbit of a planet is only a Kepler ellipse to a first approximation. When the second order terms are derived from the new theory they show that the major axis of this ellipse should precess with an angle equal to $6\pi v^2/c^2$ per revolution, v being the orbital velocity and c the velocity of light. This is exactly the amount of precession observed in the case of Mercury.

(2) The theory shows that as matter distorts the continuum, this is equivalent to the production of a variable refractive index in space in the neighbourhood of matter. Light passing close to the sun should be deflected slightly by an amount that can be calculated. Measurements made during eclipses of the sun on the apparent positions of stars close to the sun's edge com-

pletely verify the predictions.

(3) It can be proved that the frequency of a rotating or

resonating body is slowed down by a gravitational field. From this it follows that the wavelengths of spectrum lines given out by the sun are displaced slightly to the red compared with those given out on the earth, where the gravitational field is much less. The calculated displacements are very small (of the order of 0.008 A) and are just on the limit of observation. There is on the whole evidence in favour of the effect on the sun. In the very heavy stars the effect can be measured.

It can be considered that the predictions of the theory have been confirmed by observation. Thus the very remarkable theory based upon such apparently intangible ideas has been completely vindicated.

REFERENCES

[&]quot;Relativity: The Special and General Theory." A. Einstein. (1922.) "Relativity Physics." W. H. McCrea. (1933.)

APPENDIX

RECENT DETERMINATIONS OF IMPORTANT ATOMIC CONSTANTS

Within recent years a number of high precision experiments have been made in order to increase the accuracy of the experimental values for a number of fundamental atomic quantities such as the electronic mass, the electronic charge, Planck's constant, the velocity of light, etc. Frequently several of these independent quantities enter together in a particular experiment, and as each investigation involves its own percentage experimental error, a careful statistical analysis of all the experimental data must be made before the most probable values can be deduced. From the viewpoint of atomic physics, most attention has been devoted to the determination of the electronic mass and charge and indirectly it has been necessary to evaluate more accurately a number of other physical constants.

Considerable controversy has in the past existed concerning accurate values for the electronic charge e and the electronic mass at rest m_0 . With increasing precision these difficulties have now practically disappeared as will be seen from the following data compiled by Dunnington with modifications by Birge added. In order to derive the numerical values of e and m_0 from the experiments to be described, a number of auxiliary constants are required but in all of these the accuracy is such that any errors which remain do not contribute to the inconsistencies which led to the earlier controversy. The auxiliary

constants are :

TABLE I

Faraday (International coulombs per mol)
Velocity of light (cm./sec.)
Rydberg constant (cm1)
Gas constant (erg/deg. mol.)
Ratio of mass spectrographic to chemical weight
Rydberg constant (infinite nucleus)
Value adopted for Planck's constant h (erg sec.)

 $\begin{array}{l} 96506 \cdot 6 \pm 5 \\ 2 \cdot 99776 \pm 0 \cdot 00015 \times 10^{10} \\ 109677 \cdot 58 \pm 0 \cdot 01 \\ 8 \cdot 3144 \pm 0 \cdot 0004 \times 10^{7} \\ 1 \cdot 000275 \\ 109737 \cdot 30 \pm 0 \cdot 02 \\ 6 \cdot 624 \pm 0 \cdot 002 \times 10^{-27} \end{array}$

It is found that the best methods for fixing the electronic mass and charge involve first the determination of e/m_0 and then of e. We shall consider first the various e/m_0 values.

e/m_o .

The measurements on e/m_0 can be divided into two distinct groups depending upon whether the electron is bound to an atom or is free. The first group includes the precise spectroscopic methods and the second group the electric and magnetic deflection experiments. The investigation of Bearden upon the refractive index of a diamond prism for X-rays can be considered as belonging to the free electron group. The results are tabulated below.

TABLE II

Observers	Date	Method	Result × 10-
Houston	1927	Fine structure H¹—He⁴	1.7607
Kinsler and Houston .	1934	Zeeman effect	1.7571
Shane and Spedding	1935	Fine structure H ¹ —H ²	1.7582
Williams	1938	ditto	1.7580
Houston	1938	ditto	1.7593
Drinkwater, Richardson			
and Williams	1940	ditto	1.7591
Perry and Chaffee	1930	Linear acceleration	1.7610
Kirschner	1932	ditto	1.7590
Dunnington	1937	Magnetic deflection	1.7597
Shaw	1938	Crossed fields	1.7581
Bearden	1938	X-ray refraction	1.7600

Weighted mean $e/m_0 = 1.7592 \pm 0.0005 \times 10^7$ e.m.u.

The various methods employed may be described briefly as follows:—

(1) Comparison of the wavelengths of corresponding lines in the line spectra of hydrogen and helium, together with Bohr's theory of spectra, gives a precision value of e/m.

(2) When a line source is emitting in a strong magnetic field, then, in suitable cases, the lines split up into a number, from the separation of which e/m can be derived.

(3) If ν_1 and ν_2 are like components in the fine structures 22

and

of H and D lines respectively, H⁺ and D⁺ the masses of the nuclei, and F is the Faraday, then

$$e/m = (F/1 \cdot 0081)(D^{+}\nu_{1} - H^{+}\nu_{2})/D^{+}(\nu_{2} - \nu_{1})$$

$$m = H^{+}D^{+}(\nu_{2} - \nu_{1})/(D^{+}\nu_{1} - H^{+}\nu_{2}).$$

This method has very high precision.

(4) In the linear acceleration experiments, of which there are a number of variants, electrons are accelerated by a known field and the velocity attained is measured accurately.

(5) The purely magnetic deflection method employed by

Dunnington has already been described.

(6) Modern methods using crossed electric and magnetic fields are refinements and improvements on the original classical method of Thomson.

(7) From measurements of the refractive index of a diamond prism for X-rays of known wavelength, e/m can be found in terms of the known grating space of the diamond crystal.

It is clear from the data in Table II that the different values have a random distribution so that within the present limits of measurement the value of e/m_0 cannot be considered to be affected by the degree of binding of the electron.

е

The electronic charge e can only be measured directly with precision by the oil-drop method, in which the velocity of a drop in a gravitational field and then in a combined gravitational and electrical field is measured. There are, however, at least ten other ways by which the electronic charge can be indirectly evaluated. In these experiments various combinations of e, m_0 , and h occur, and by combining these where necessary with the known values of Rydberg's constant or the velocity of light c, appropriate methods of algebraic elimination can be applied and from each a value of e can be derived. The results obtained by these methods are tabulated in Table III.

The methods may be briefly described in turn as follows:-

(1) The wavelength of a particular X-radiation is measured with a ruled grating and then with a calcite crystal. From the derived grating space of the latter the charge on the electron can be calculated.

TABLE III

Method	Combination measured	e×1010 e.s.u.
Ruled grating		4.8025
X-ray continuum limit	h/e	4.8026
Ionisation potentials	h/e	4.8090
Wien's radiation constant	h/e	4.8145
Stefan-Boltzmann radiation constant .	e/hŧ	4.8168
Electron diffraction (crystal)	$(h/e)(e/m_0)$	4.7964
Electron diffraction (film)	h/m_n	4.7972
Compton effect	h/m_0	4.7956
Specific charge	e/m _o	4.7963
X-ray photo-effect	$(e/m_0)(e/h)$	4.7953
Oil drop method	ě	4.8036
Oil drop method	e	4.802

Weighted mean $e=4.8025\pm0.0004\times10^{-10}$ e.s.u.

(2) In this method the minimum voltage at which electrons can produce X-rays of specified wavelength is determined.

(3) Ionisation potentials are measured by the method of

electronic impact.

(4) Wien's radiation constant is derived from heat radiation observations.

(5) The Stefan-Boltzmann constant is obtained by measuring the net radiation per square centimetre from a heat source at known temperature.

(6) Electron diffraction observations with a crystal give the de Broglie wavelengths of the electrons used in terms of the

crystal grating space.

(7) The diffraction of electrons from surface films, when

examined, leads to a value for h/m_0 .

(8) The same quantity is obtained from Compton effect studies on the scattering of X-rays, the effect of binding being taken into account.

(9) The methods for evaluating the specific charge have

already been discussed.

(10) X-ray photo-electrons are ejected from a thin film with radiation of known frequency and the electron energies measured using a magnetic deflection method.

(11) The largest error in the oil-drop method lies in the value of the coefficient of viscosity of air. This has recently

been re-determined and has led to a modification of the earlier value reported for the charge by Millikan.

The best value for the mass of the electron, derived from the charge and the ratio of mass to charge is

$$m_0 = 9.1066 \times 10^{-28}$$
 gms.

To the quantities already enumerated three more can be added. According to Millikan (1938) the best value for Avogadro's number is

$$N=6.023\pm0.001\times10^{23}$$
 per mol.

With this the mass of the hydrogen atom becomes

$$M = 1.674 \times 10^{-24}$$
 gms.

The third quantity is the mass of the neutron. This is best measured by disintegrating heavy hydrogen with gamma-radiation of known wavelength. The most recent value found for the mass (Chadwick, Feather and Bretscher, 1939) in terms of oxygen=mass 16 is

neutron mass=1.00895.

REFERENCE

Review of Modern Physics. April, 1939. Dunnington. Reports on Progress in Physics. 8.1941. Birge.

APPENDIX II

NUCLEAR FISSION AND ATOMIC ENERGY

Explanatory

In Chapter XVIII the nuclear fission of uranium was briefly mentioned and it was hinted that this might constitute a vast source of nuclear energy. Research in this field, culminating in the atomic bomb explosions of 1945, was being intensively pursued secretly throughout 1939–45, but it is only recently that the secrecy ban has been lifted and an account of this work can thus be included in this reprint as an appendix.

Nuclear Fission

In accordance with the Einstein principle of the equivalence of energy and mass ($E=mc^2$ where c is the velocity of light) it follows that a kilogramme of matter, if completely transformed into heat energy, could convert over 30 million tons of water into steam. This conversion is the source of stellar energy of radiation, and apart from separate individual atomic disintegrations naturally associated with cosmic rays and radioactivity, no such energy conversions normally take place on the earth.

The practical application of nuclear disintegration processes to both the production of an atomic bomb and to the future probable employ as a source of power, owes its origin to the utilization of the nuclear fission of uranium, first demonstrated by Hahn and Strassmann in 1938, the mechanism of which will now be discussed. A comprehensive theory of the process has been developed by Bohr and Wheeler with the following general conclusions.

If a neutron be captured by a heavy nucleus then there will be either (1) photon radiation, or (2) neutron emission, or (3) fission. There is competition between these processes, depending largely upon the state of stability of the nucleus. If the energy of capture is transformed into deformation of the nuclear "drop," as postulated by Bohr, then under suitable conditions the "drop" deforms so far as to split in two, since excitation of the nucleus by absorption of the neutron leads to oscillations like those produced in a water drop by surface tension. Some very heavy nuclei are near the limit of stable constitution, and it does not require great energy to initiate fission.

Experiments made by separating the fission products arising from the neutron bombardment of uranium established that the latter splits up into two nuclei of roughly equal mass, with a great liberation of energy. The two new, charged nuclei. being formed within the small distance of a nuclear radius, are violently repelled with great energy. There is an energy gain of some 200 million electron volts at the expense of mass consumed in the process. In 1939 Joliot, Halban and Kowarski showed that the fission of one uranium atom by one neutron is accompanied by the simultaneous emission of three other neutrons. It is this latter fact which permits of application of fission to practical purposes. For it enables initiation of a chain reaction in a sufficiently large mass of material, resulting in a rapid geometric increase in fissions, once the first fission has started, in a manner recalling the formation of a cosmic ray cascade shower.

Bohr and Wheeler showed that a rare isotope of uranium was involved in the observed fission. There are three isotopes in uranium masses 238, 235, 234 (abundance ratios 1:1/140:1/17,000). These investigators showed that the binding energy of a neutron in a heavy nucleus of even charge is less when the mass is odd than when it is even. This points to 235 as the responsible atom and this was confirmed by energy measurements. It was also shown that 238 could undergo fission too but only when bombarded by high energy neutrons, whilst 235 required only low energy (thermal) neutrons to initiate fission. This prediction of the theory was confirmed by extracting a small amount of 235, using a large mass spectrograph as the instrument for isotope separation.

The mechanism of fission of 235 by neutron absorption can thus be represented by the reaction

$$_{92}\mathrm{U}^{235}+_{0}n^{1}\rightarrow_{92}\mathrm{U}^{236}$$
 $\left\{ \begin{array}{c} \longrightarrow & \mathrm{neutrons} \\ \longrightarrow & (a) \text{ fission product} \\ \longrightarrow & (b) \text{ fission product} \end{array} \right.$

The U^{236} is unstable and breaks up in a short time period estimated to be 10^{-15} second, emitting *high* speed neutrons.

The two fission products have variable masses, one being in the range 124-134 and the other in the range 90-100. Both products are unstable and decay by successive beta emissions with periods ranging from fractions of a second to over one year.

The Chain Reaction

Theory of the reaction showed that if a chain reaction could at all be initiated, then for equal weights, the energy produced would be many millions of times that of chemically burning fuels. Two conditions are necessary for onset of a chain reaction: (a) the mass of reacting material must be large enough to absorb and conserve the neutrons which would otherwise escape, leading to collapse of the chain of collisions; (b) the purity must be high, that is, the quantity of non-active material capable of absorbing (and wasting) neutrons, must be strictly reduced. With U235 the critical mass is shown by calculation to be only a few kilograms (less than 100). But the U235 chain is normally prevented through absorption by the 140 times more abundant U238, which is in this sense an impurity. If U235 can be isolated reasonably free from admixture, the critical mass will spontaneously lead to a violent chain reaction, exploding with a power calculated to be of the order of that of 8,000 tons of TNT. This fission explosion only occurs when the critical mass is reached, for in a smaller volume of metal less than one neutron is conserved per fission, as there is escape at the surface and the chain fails to develop. For a sphere of metal the neutron escape depends on the surface, i.e. on the square of the radius, but the absorption depends on the volume, i.e. the cube of the radius. Clearly as the radius is increased, from a small value, a critical size is reached and above this more neutrons are produced than are lost, and the explosion develops rapidly.

Quantities less than the critical amount are completely stable and quantities exceeding this spontaneously explode. A mass of metal exceeding the critical mass cannot therefore exist.

To produce the explosion it is only necessary to bring inti-

mately together mechanically two quite safe pieces of 235 metal each of mass somewhat greater than half the critical value. This mechanical operation constitutes the act of detonation. No other mechanism for artificially inducing the explosion is required, for there are always produced in any mass of metal a number of free neutrons by incident cosmic radiation, and these suffice to start the explosion.

Nature of the explosion

If an appreciable fraction of the atoms in a mass of 235 metal break up rapidly by fission, the energy thus liberated leads to a violent rise in temperature, which might be of the order of 10,000,000° C. and will be accompanied by a pressure of many millions of atmospheres. Radiation, being proportional to the fourth power of the temperature, will be intense and the associated radiation pressure will be correspondingly high. There is also intense gamma radiation and a great cloud of radioactive matter forms.

Under such conditions it is clear that the disintegrating material will itself expand at great speed. As the expansion leads automatically to a fall in density, neutron absorption will diminish and the chain reaction will quickly collapse. It is essential that the fission reaction chain will develop so rapidly that a considerable portion of the material reacts before dispersal. The life-time of the reaction is small enough and the neutrons are sufficiently fast to ensure that this important condition is actually fulfilled.

The life-time of the total explosion itself is very small and is determined by the rate of expansion which ultimately leads to cessation of fission. All the 235 is not converted and the calculations which have been made on the energy liberated take into account only the fraction consumed.

Separation of U235

The production of several kilogrammes of separated U²³⁵ represented a major national effort in the U.S.A. The processes tried included diffusion, distillation, centrifuging, thermal diffusion, chemical exchange and separation by mass spectrograph, using a 2,000 ton magnet. The major separation was

achieved by gaseous diffusion using a gaseous compound possibly like UF₆. The two isotopic compounds $U^{235}F_6$, $U^{238}F_6$ have masses 349, 352, hence a single diffusion through

a porous barrier enriches U235F₆ by a factor $a = \sqrt{\frac{352}{349}} = 1.0043$.

It is such small ratios which lead to the installation of a vast plant before practical separation can be achieved. Many thousands of stages are required and the initial quantity of gas needed is some 100,000 times the quantity ultimately obtained, so that 100 kilogrammes separated require 10,000 tons of uranium. Thousands of pumps and many acres of diffusing barrier were necessary. The cost was of the order of a hundred million pounds. The achievement is unique in the history of science.

Sufficient U²³⁵ was obtained for the manufacture of a small number of atomic bombs. The project was highly speculative since no preliminary small-scale test bomb can be constructed. The final production and detonation of the bombs confirmed the whole theoretical prediction.

The Pile

Experiments showed that at certain neutron speeds U²³⁸ had a large probability of capture, but instead of fission, U²³⁹ is formed. This is unstable (beta active) and on decay leads to a transuranic element, i.e. one of atomic number greater than 92. This element 93Np²³⁹ has been named neptunium and the reaction is

$$_{92}\mathrm{U}^{238}+_{0}n^{1}\rightarrow_{92}\mathrm{U}^{239}\rightarrow_{93}\mathrm{Np}^{239}+e.$$

Neptunium is beta active and decays to another new element 94 Pu 239 , plutonium, thus

$$_{93}$$
Np²³⁹ \rightarrow_{94} Pu²³⁹+e.

The importance of this reaction is that plutonium is easily separated *chemically* from the uranium and in accordance with the Bohr-Wheeler theory its fission properties should closely resemble U²³⁵.

Now materials of low atomic weight (e.g. carbon, but preferably deuterium) can slow down fast neutrons. If ordinary uranium be mixed with such a "moderator" any fast neutrons from U²³⁵ fissions that may be initiated by cosmic rays, etc.,

will be slowed down and then absorbed by U^{238} to form Pu^{239} ultimately.

A large complex lattice of lumps of uranium distributed within a moderator is called a "pile." By such a device the neutron velocities can be slowed down to the required values for producing fissions, and the U²³⁸ atoms do not effectively destroy the reaction chain (neutrons of intermediate energies have a resonance reaction with 238).

Such a large pile can be maintained self-sustaining and thus become a power plant. Thus the production of power requires a slow-neutron induced fission chain in a pile, whilst an atomic bomb requires a fast-neutron chain in U^{235} or Pu^{239} .

In 1942 a small pile was produced using pure graphite as moderator. As the pile was built up the critical dimension for the particular concentration of 235 used was reached (several tons) and additional mass led to a self-sustained reaction. The first pile contained 6 tons of uranium and produced the low power of 200 watts, but this was sufficient to confirm the theory. The reaction is controllable by the insertion of strips of absorbing metal (cadmium). To produce 1 kilogramme of Pu²³⁹ per day the pile must operate at about 1 million kilowatts, from which it is seen that an effective Pu²³⁹ producing pile must contain thousands of tons of uranium metal and represents a vast engineering project.

The pile warms up and the cooling problem is a great one. Furthermore the uranium must be protected from corrosion and oxidation. To add to the difficulties the neutron radiation from the pile is so great and dangerous that *all* operations within this vast structure must be automatically controlled from a considerable distance away. No operators can approach a pile, which has to be enclosed in thick concrete walls.

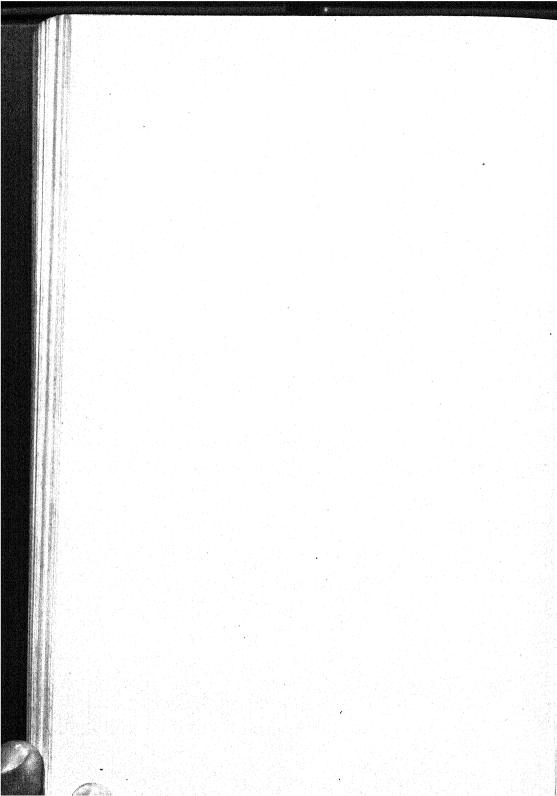
The individual piles, of which a number have been constructed, consume only a few grammes per day of uranium.

An opening at the top of the pile lets out an intense uniform beam of neutrons, which serve as a source for further experiments.

The piles now in operation function principally as producers of Pu²³⁹ and other radioactive materials. The energy liberated is not at present available for practical purposes since it is the low efficiency energy in the heated up

water-cooling system, and is allowed to go to waste. The technological problems of running a pile at high temperature have not yet been solved. When they are, the piles may become practical sources of power on a great scale. The pile is permitted to run for some time and pieces of uranium are subsequently removed and the plutonium extracted chemically. The recovered uranium is used again, and by this means a stock of plutonium is slowly accumulated.

REFERENCE
Atomic Energy. H.M. Stationery Office, 1945.



INDEX

Aberration of light, 321 a-particleabsorption, 235 charge, 233 detection, 230 E/M, 232 fine structure, 239 Gamow's theory, 296 properties, 246 ranges, 237 scattering, 241 Artificial radioactivity, 307 Aston's mass spectrograph, 62 Atomic constants-Avogadro's number, 49 e, 41 electron mass, 48 neutron mass, 299

Balmer series, 90 \$\beta\$-rays—
absorption, 262
continua, 260
disintegration, 258
e/m, 250
ionising power, 263
mass increase, 253
origin, 256
scattering, 261
spectra, 255
Bohr hydrogen theory, 90
Brownian motion, 49
Brush discharge, 20

Cathode rays, 31 Cathode sputtering, 33 Cold electronic emission, 111 Compton effect, 161 Cosmic radiationabsorption, 274 azimuth effect, 281 geographical distribution, 276 geomagnetic effect, 280 nature of, 277 showers, 287 time variation, 277 Crystal analysisionisation spectrometer, 173 lattices, 168 Laue method, 172

Crystal analysis—
powder method, 173
rotation method, 173
Crystal diffraction, 170
Crystal structure—
complex, 178
Fourier analysis, 180
KCl, 175
metals, 183
NaCl, 175
organic, 185
Crystal texture, 186
Cyclotron, 305

Deuterium, 70

Einstein photo-electric law, 119 e/m— Thomson, 45 Dunnington, 47 Electric image force, 109 Electrical conductivity, gases, 2 Electrical conductivity, metals, 108 Electrodeless discharge, 36 Electroncharge, 41 diffraction, 189 distribution states, 199 dualistic nature, 192 elastic scattering, 205 mass, 48 Schroedinger equation, 197 wavelength, 188 Electron theory of metals, 104 y-raysabsorption, 267 emission, 269 scattering, 267 wavelengths, 265 Gas dischargemechanism, 29 potential variations, 27 reduced pressures, 26 Geiger counter, 232 Geissler tube, 34

Helium spectrum, 131 High voltages, 303 Hollow cathode discharge, 35 Hydrogen atomic mass, 53

Interatomic forces, 182 Ionisation by collision, 14 Ionsage effect, 10 condensation properties, 21 condensation theory, 22 impurities, 13 mass and mobility, 11 mobility, 5 nature of, 8 pressure effect, 13 recombination, 4 temperature effect, 13 Isotope separation, 71 Isotope spectra, 149 Isotopes, radioactive, 222 Isotopic constitution, 65

Light counters, 129 Lorentz-Fitzgerald contraction, 323

Mass defect, 68
Mass spectra, parabolas, 54
Mass spectrograph, Aston, 59
Mass spectrograph, Bainbridge, 64
Meson—
instability, 291
properties, 292
Yukawa theory, 290
Metals, electron theory, 104
Metals, electron states, 106
Michelson-Morley experiment, 318
Moseley's law, 155

Neutrino, 260 Neutron, 299 Nuclear disintegration a-particles, 298 Bohr's theory, 301 fission, 311 stellar energy, 312 Nucleusartificial disintegration, 298 chemistry of, 308 energy levels, 270 fission of, 311 Gamow's theory, 296 isomerism, 311 spin, 147 spin distributions, 315 transmutations, 298 transitions, 317

Oxide-coated cathode, 103

Packing fraction, 68 Paschen's sparking law, 19 Pauli exclusion principle, 143 Periodic table, 144

Photo cellsconductivity, 129 emission, 126 light counters, 129 rectifier, 128 Photo effect-Einstein's law, 119 fatigue, 123 inner, 125 outer, 125 Photo electronse/m, 141 effect of wavelength, 121 non-metallic, 124 velocities, 116 Positive rays, 54 Positron, 282 Positron mass, 284 Positron, nature of, 285

Quantum theory—heat radiation, 75 specific heat, 82 spectra, 88

Radiation theory, classical, 75 Radiation theory, quantum, 79 Radioactivity. active deposits, 218 artificial, 307 decay, 212 disintegration, 213 emanations, 216 equilibria, 214 haloes, 247 isotopes, 222 series, 219 Radium, 225 Radon, 217 Ramsauer effect, 203 Rayleigh-Jeans formula, 75 Relativityfour dimensional continuum, 332 General theory, 334 mass variation, 331 Special theory, 325 velocity of light, 330 Schottky effect, 111

Spark discharge, 18
Specific heat theories—
classical, 79
Debye, 84
Einstein, 82
Spectra—
controlled excitation, 208
elliptic orbits, 134
hyperfine structure, 147

intensity rules, 143

Shot effect, 112

Spectra—
interval rules, 143
isotope effect, 149
many electron, 142
one electron, 136
series, 133
spining electron, 139

Thermionic emission electron velocities, 98 impurity effect, 102 potential effect, 100 surface adsorption, 103 temperature effect, 96

Uncertainty Principle, 194

Vacua, 37

Wave mechanics, 188 Whole number rule, 66 Wilson cloud chamber, 23

X-rays—absorption, 159
continuous spectrum, 158
diffraction, 170
emission spectrum, 152
energy levels, 157
interference, 163
production of, 151
reflection, 165
refraction, 163